

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5781925号  
(P5781925)

(45) 発行日 平成27年9月24日(2015.9.24)

(24) 登録日 平成27年7月24日(2015.7.24)

(51) Int. Cl. F 1  
**G 0 6 F 3 / 0 6 (2006.01)**  
 G 0 6 F 3 / 0 6 3 0 1 E  
 G 0 6 F 3 / 0 6 3 0 1 Z

請求項の数 11 (全 38 頁)

|               |                              |           |   |
|---------------|------------------------------|-----------|---|
| (21) 出願番号     | 特願2011-510141 (P2011-510141) | (73) 特許権者 | 000005108<br>株式会社日立製作所<br>東京都千代田区丸の内一丁目6番6号     |
| (86) (22) 出願日 | 平成21年4月23日(2009.4.23)        | (74) 代理人  | 100093861<br>弁理士 大賀 真司                          |
| (86) 国際出願番号   | PCT/JP2009/058533            | (74) 代理人  | 100129218<br>弁理士 百本 宏之                          |
| (87) 国際公開番号   | W02010/122674                | (72) 発明者  | 里山 愛<br>神奈川県横浜市戸塚区吉田町292番地<br>株式会社日立製作所 横浜研究所内  |
| (87) 国際公開日    | 平成22年10月28日(2010.10.28)      | (72) 発明者  | 江口 賢哲<br>神奈川県横浜市戸塚区吉田町292番地<br>株式会社日立製作所 横浜研究所内 |
| 審査請求日         | 平成23年8月2日(2011.8.2)          |           |   |
| 審判番号          | 不服2014-7830 (P2014-7830/J1)  |           |   |
| 審判請求日         | 平成26年4月25日(2014.4.25)        |           |   |

最終頁に続く

(54) 【発明の名称】 計算機システム及びその制御方法

(57) 【特許請求の範囲】

【請求項1】

記憶媒体を有する記憶デバイスと、  
 上位装置からのアクセスが許容される仮想ボリュームと、  
 プールを構成し、それぞれ記憶領域を有する複数の論理ボリュームと、  
 前記プールの前記複数の論理ボリュームから前記仮想ボリュームに対する前記記憶領域の割り当て単位での割り当てを制御するコントローラと、を備え、  
 前記コントローラは、  
 前記上位装置から前記仮想ボリュームの前記記憶領域の割り当て単位の割り当て前の領域にライトアクセスがあった場合に、前記仮想ボリュームの当該領域に対して、前記プールの前記複数の論理ボリュームのいずれかの論理ボリュームから前記仮想ボリュームに割り当てられていない少なくとも1つの前記記憶領域の割り当て単位を割り当て、  
 前記プールの前記複数の論理ボリュームのそれぞれの容量と、前記複数の論理ボリュームそれぞれが有する前記記憶領域のうち前記仮想ボリュームに対する割り当てに使用されている前記記憶領域の割り当て単位の数による容量と、を含むプール情報を管理し、  
 前記複数の論理ボリュームのそれぞれについて、当該論理ボリュームの容量のうち、前記仮想ボリュームへの割り当てに使用されている前記記憶領域の割り当て単位の割り当て数による容量の割合である使用率を、前記プール情報に基づいて決定することにより、当該決定した使用率が前記複数の論理ボリューム間で不均衡になる状況を検出し、そして、  
 前記検出に応じて、前記複数の論理ボリューム間において前記使用率が均等になるよう

10

20

に、前記複数の論理ボリュームのうち少なくとも1つの論理ボリュームの前記記憶領域の割り当て単位の前記仮想ボリュームに対する割り当てを、他の前記論理ボリュームの記憶領域の割り当て単位の当該仮想ボリュームに対する割り当てに変更する割り当て変更処理を実行する、計算機システム。

【請求項2】

前記コントローラは、前記複数の論理ボリュームに属する第1の論理ボリュームと第2の論理ボリュームについて、前記第1論理ボリュームから前記仮想ボリュームに割り当てられた前記記憶領域の第1の割り当て単位の当該割り当てを、前記第2の論理ボリュームの記憶領域の第2の割り当て単位の前記仮想ボリュームへの割り当てに変更しようとする際、前記仮想ボリュームの前記第1の割り当て単位が割り当てられた領域に、前記上位装置からのアクセスがあった場合には、前記仮想ボリュームに対する当該割り当ての変更を中断し、

10

前記仮想ボリュームに割り当てられた前記記憶領域の割り当て単位が、固定パターンデータを格納している場合、当該仮想ボリュームに割り当てられていた前記記憶領域の属性が前記割り当ての未実行の属性に変更されるようにするとともに、

当該仮想ボリュームの領域のうち、前記固定パターンデータを格納した前記割り当て単位が割り当てられていた領域に、前記固定パターンデータを格納した専用領域を割り当てるように構成されている、請求項1記載の計算機システム。

【請求項3】

前記コントローラは、前記プールに新たな論理ボリュームを追加したことを前記不均衡が発生する状況として検出する、請求項1記載の計算機システム。

20

【請求項4】

前記コントローラは、前記複数の論理ボリュームのうち、前記使用率が所定値より高い前記論理ボリュームの記憶領域の割り当て単位の前記仮想ボリュームに対する割り当てを解放し、その割り当てを前記使用率が前記所定値より低い前記論理ボリュームの記憶領域の割り当て単位に変更する、請求項1記載の計算機システム。

【請求項5】

前記コントローラは、前記仮想ボリュームに対する割り当てが解放された前記記憶領域の割り当て単位の属性を割り当て済みから未割り当てに変更する、請求項4記載の計算機システム。

30

【請求項6】

前記コントローラは、前記プールから前記複数の論理ボリュームのうちの一部の論理ボリュームの削除を前記不均衡が発生する状況として検出する、請求項1記載の計算機システム。

【請求項7】

前記コントローラは、前記削除される論理ボリュームの記憶領域の割り当て単位の前記仮想ボリュームに対する割り当てを、前記複数の論理ボリュームのうちの前記プールに削除されることなく維持され、前記削除される論理ボリュームとは別の論理ボリュームの記憶領域の割り当て単位の前記仮想ボリュームへの割り当てに変更する際、前記プールに維持される複数の論理ボリューム間で、前記仮想ボリュームへ割り当てられる前記記憶領域の割り当て単位の割り当てを変更する、請求項6記載の計算機システム。

40

【請求項8】

前記コントローラは、前記複数の論理ボリュームに属する第1の論理ボリュームと第2の論理ボリュームについて、前記第1論理ボリュームから前記仮想ボリュームに割り当てられた前記記憶領域の第1の割り当て単位の当該割り当てを、前記第2の論理ボリュームの記憶領域の第2の割り当て単位の前記仮想ボリュームへの割り当てに変更しようとする際、前記仮想ボリュームの前記第1の割り当て単位が割り当てられた領域に、前記上位装置からのアクセスがあった場合には、前記仮想ボリュームに対する当該割り当ての変更を中断する、請求項1記載の計算機システム。

【請求項9】

50

前記コントローラは、前記仮想ボリュームに割り当てられた前記記憶領域の割り当て単位が、固定パターンデータを格納している場合、当該仮想ボリュームに割り当てられていた前記記憶領域の属性が前記割り当ての未実行の属性に変更されるようにするとともに、当該仮想ボリュームの領域のうち、前記固定パターンデータを格納した前記割り当て単位が割り当てられていた領域に、前記固定パターンデータを格納した専用領域を割り当てるように構成されている、請求項 1 記載の計算機システム。

【請求項 10】

前記固定パターンデータは、全 0 データからなるものである、請求項 9 記載の計算機システム。

【請求項 11】

前記コントローラは、  
前記複数の論理ボリュームのそれぞれについて、前記仮想ボリュームに割り当てられた記憶領域の割り当て単位の数、及び（前記複数の論理ボリュームのそれぞれの容量 / 前記複数の論理ボリューム全体の容量）を計算し、

当該計算の結果に基づいて、前記割り当て変更処理を実行する、請求項 1 記載の計算機システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、計算機システムに係わり、特に、ホスト装置に対して動的に記憶容量を割り当てる計算機システムに関するものである。

【背景技術】

【0002】

従来から、ホスト装置に対して大規模なデータストレージサービスを提供する計算機システムが存在する。このシステムは、ホスト装置と、ホスト装置が接続するストレージ装置と、ストレージ装置の管理装置と、を備えたものとして知られている。

ストレージ装置は、複数のハードディスクを RAID (Redundant Array of Independent / Inexpensive Disks) 方式で管理する。そして、多数のハードディスクが有する物理的な記憶領域を論理化し、これを論理ボリュームとしてホスト装置に提供する。ホスト装置は論理ボリュームにアクセスしてデータのリード・ライトを要求する。

この種の論理化技術の一つにシン・プロビジョニングと称されるものがある。ストレージ装置は、物理的な記憶領域を有さない仮想的なボリュームをホスト装置に対して設定する。ホスト装置が仮想的なボリュームにライトアクセスにすることに合わせて、ストレージ装置は、仮想ボリュームに対して記憶領域を順次割り当てていく。

したがって、この技術は、論理ボリュームに対して当初から大容量な記憶領域を割り当てる方式に比べて、ストレージ資源を有効に活用できる点で効果的である。この技術は、特願 2003 - 15915 号公報、及び、特開 2006 - 338341 号公報に記載されている。

仮想ボリュームに記憶領域を提供するための手段は、それぞれ実の記憶領域が割り当てられた複数の論理ボリュームをプールと呼ばれる論理グループに纏めたものから構成される。複数の論理ボリュームはそれぞれプールボリュームと呼ばれる。仮想ボリュームにホスト装置からライトアクセスがあると、仮想ボリュームのアクセス先に、プールボリュームの対応領域の記憶領域を割り当てる。

ストレージ装置は、仮想ボリュームに対するホスト装置からのアクセス領域を、プールボリュームに記憶容量が割り当てられた領域に対応させることにより、ホスト装置からのライトデータを保存することができる。

ストレージ装置は、複数のプールボリュームへの記憶領域の割り当てが均等になるようにするが、長い時間を経過してゆく過程で、この均等が崩れてくる問題がある。そこで、特開 2008 - 234158 号は、複数のプールボリューム間で記憶領域を移動させて、

10

20

30

40

50

記憶領域の容量が複数のプールボリュームについて均等になることを開示している。

なお、米国特許第6857059号公報には、仮想ボリュームの未割当て領域へリードアクセスが来た場合、ストレージ装置は、プログラムによって全部が“0”からなる固定パターンを作成してホスト装置へ応答すること、仮想ボリュームの未割当て領域へライトが来た場合、記憶領域を割り当てることが記載されている。

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特願2003-015915号公報

【特許文献2】特開2006-338341号公報

【特許文献3】特開2008-234158号公報

【特許文献4】米国特許第6857059号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

特開2008-234158号は、外部操作によって、記憶領域を移動させることを開示しているものの、複数のプールボリュームについて、記憶容量の均等化が迅速、かつ確実に実行されることまでを提案するものではなかった。例えば、プールボリュームから固定パターンデータを破棄すると、それだけで、複数のプールボリュームについての記憶容量の均等化が損なわれることにもなってしまう。

そこで、本発明は、仮想ボリュームに記憶領域を割り当てる複数の論理ボリュームに対して、記憶容量の均等化が迅速、かつ確実に実行される計算機システムを提供することを目的とするものである。

さらに、本発明は、固定パターンデータを削除することなどのように、論理ボリュームのデータ量に変化が避けられない場合にも、仮想ボリュームに記憶領域を割り当てる複数の論理ボリュームに対して、記憶容量の均等化が維持される計算機システムを提供することを目的とするものである。

【課題を解決するための手段】

【0005】

この目的を達成するために、本発明は、上位装置からのアクセスに対して、仮想ボリュームに記憶領域を動的に割り当てる運用を実行している過程で、複数の論理ボリューム間で記憶容量のバランスが崩れることを検出し、その後、記憶領域を複数の論理ボリューム間で移動させて、記憶容量のバランスを維持するようにした計算機システムであることを特徴とするものである。

本発明は、さらに、論理ボリュームから固定パターンデータの削除を行いながら、複数の論理ボリューム間で記憶容量のバランスを図るようにした計算機システムであることを特徴とするものである。

【発明の効果】

【0006】

本発明によれば、仮想ボリュームに記憶領域を動的に割り当てる複数の論理ボリュームに対して、それぞれの記憶容量の均等化が迅速、かつ確実に実行される計算機システムを提供することができる。

さらに、本発明によれば、固定パターンデータを削除することなどのように、論理ボリュームのデータ量に変化が避けられない場合にも、仮想ボリュームに記憶領域を割り当てる複数の論理ボリュームに対して、記憶容量の均等化が維持される計算機システムを提供することができる。

【図面の簡単な説明】

【0007】

図1は、本発明に係わる計算機システムの第1の例を示すハードウェアブロック図である。

10

20

30

40

50

図 2 は、図 1 に示した計算機システムにおいて、ストレージ装置が行う記憶領域の動的割り当ての動作を示す機能ブロック図である。

図 3 は、ストレージ装置のメモリブロック図である。

図 4 は、VDEV 管理情報を説明するブロック図である。

図 5 は、LDEV 管理情報を説明するブロック図である。

図 6 は、アドレス管理テーブルを説明するブロック図である。

図 7 は、POOL 管理情報のブロック図である。

図 8 は、VVOL - DIR 及び PSCB のブロック図である。

図 9 は、計算機システムの第 2 の形態を示すハードウェアブロック図である。

図 10 A は図 1 に示すシステムにおける、ターゲットデバイスと、LDEV と、PDEV とのマッピングの関係を示すブロック図であり、図 10 B は図 9 のシステムの同ブロック図である。

10

図 11 に、1 つの仮想ボリュームと複数のプールボリュームとの間で、均等に、ページが割り当てられている様子を示すブロック図である。

図 12 は、複数のプールボリューム間でのページのリバランスの過程を示す仮想ボリューム及びプールボリュームのブロック図である。

図 13 は、複数のプールボリューム間で、プールボリュームの容量比を勘案してリバランスを行う様子を示す、仮想ボリューム及びプールボリュームのブロック図である。

図 14 は、リバランスの契機にしたがって、ストレージ装置がリバランスの処理を行うことを説明するフローチャートである。

20

図 15 はリバランスを示すフローチャートである。

図 16 A はプールを管理するテーブルであり、図 16 B はプールボリュームを管理するテーブルであり、図 16 C は、フリーキュー単位を管理するテーブルである。

図 17 は、ホスト装置から仮想ボリュームへのライトアクセスに対して、プールに属するプールボリュームから仮想ボリュームにページを割り当てる順序を示したブロック図である。

図 18 は、この割り当て順序を管理するテーブルである。

図 19 は、複数のプールボリューム間でマイグレーションされるページ数の算出の処理を示すフローチャートである。

図 20 は、算出されたページ数とページの移動の様子を示す、仮想ボリュームとプールボリュームとのブロック図である。

30

図 21 は、図 20 に示すページの移動形態を実現するための管理テーブルである。

図 22 は、ページ割り当てバランス管理テーブルである。

図 23 は、マイグレーション処理の詳細を示すフローチャートである。

図 24 は固定データパターンの破棄を伴う、マイグレーション処理の詳細を示すフローチャートである。

図 25 は、ページマイグレーションに対してホスト装置から I/O 処理を優先させることを説明するフローチャートである。

図 26 は、図 24 の処理を実現するためのプール及びキャッシュメモリのブロック図である。

40

図 27 A は、仮想ボリュームの未使用領域がプールボリュームの 0 データ専用領域に割り当たっている様子を示す仮想ボリュームとプールボリュームとのブロック図であり、図 27 B は仮想ボリュームにプールボリュームの未使用領域が割り当てられたことを示すブロック図であり、図 27 C はプールボリュームの割り当てページが全 0 データである場合の処理内容を説明するブロック図である。

図 28 は、0 データ専用領域へのリード処理を示すフローチャートである。

図 29 は、仮想ボリュームの未割当領域にホスト装置からライト要求が発行された場合の処理を説明するフローチャートである。

図 30 A は、追加プールボリュームに 0 データ専用領域を設定している様子説明する仮想ボリュームとプールボリュームとのブロック図であり、図 30 B は、削除されるプー

50

ルボリュームに0データ専用領域が割り当てられている様子を説明するブロック図である。

図31は図26の変形例に係るブロック図である。

図32は、複数のプールボリューム間でページマイグレーションが行われている様子を説明するブロック図であり、図33の前提となるものである。

図33はページのマイグレーションを実現するフリーキュー処理を説明するブロック図である。

図34は、図35の前提となるブロック図である。

図35は、マイグレーションの中断が発生した場合でのフリーキュー処理を説明するブロック図である。

図36は、マイグレーション進捗管理のための管理テーブルである

図37は、リバランスを考慮したプールボリュームのシュリンク処理を説明するフローチャートである。

図38は、GUIによってユーザに提供されるシュリンク情報の例である。

図39はシュリンクの予想結果の出力例である。

図40はストレージ装置がシュリンクするプールボリュームを決定するための処理の詳細を示したフローチャートである。

図41及び図42は、プールボリュームの削減を伴ったリバランスを説明するためのブロック図である。

図43は、図9に示すストレージ装置が複数のクラスタを備えていることを説明するハードウェアブロック図である。

図44は、仮想ボリュームの属性に、ホスト装置からのアクセスに対する応答の高速性を示す「性能」を追加したことを説明する仮想ボリュームとプールボリュームとのブロック図である。

【発明を実施するための形態】

【0008】

次に本発明の実施形態について説明する。図1は、本発明に係わる計算機システムの第1の例を示すハードウェアブロック図である。計算機システムは、ホスト計算機10と、管理装置20と、これらが接続するストレージ装置30と、を備えている。ストレージ装置は、ストレージシステム、あるいはストレージサブシステムとも呼ばれている。

ホスト計算機10は、ストレージ装置30の論理的な記憶資源にアクセスする。管理装置20は、ストレージ装置30の記憶領域の構成を管理する。ストレージ装置30は、物理デバイス34に設定された記憶領域にデータを格納する。

ホスト計算機10は、入力手段110、出力手段120、CPU130、メモリ140、ディスクアダプタ150、ネットワークアダプタ160及びディスクドライブ170を備える。入力手段110は、ホスト計算機10を操作する管理者等から入力を受け付ける手段である。入力手段110は、例えば、キーボードで構成される。出力手段120は、ホスト計算機10の状態や設定項目を表示する手段である。出力手段120は、例えばディスプレイ装置で構成される。

CPU130は、ディスクドライブ170に格納されているプログラムをメモリ140に読み込んで、そのプログラムに規定された処理を実行する。メモリ140は、例えばRAM等で構成され、プログラムやデータ等を格納する。

ディスクアダプタ150は、ストレージ装置30とストレージネットワーク50を介して接続し、ストレージ装置30にデータを送受信する。ストレージネットワーク50は、データ転送に適したプロトコル(例えば、Fibre Channel)で構成される。

ネットワークアダプタ160は、ストレージシステム管理装置20又はストレージ装置30と管理ネットワーク40を介してデータを送受信する。管理ネットワーク40は、例えばEthernetで構成される。

ディスクドライブ170は、例えばハードディスク装置で構成され、データやプログラムを格納する。

10

20

30

40

50

管理装置 20 は、入力手段 210、出力手段 220、CPU 230、メモリ 240、ネットワークアダプタ 250 及びディスクドライブ 260 を備える。

入力手段 210 は、管理装置 20 を操作する管理者等の入力を受け付ける手段である。入力手段 210 は、例えばキーボードで構成される。出力手段 220 は、管理装置 20 の状態や設定項目を表示する手段である。出力手段 220 は、例えばディスプレイ装置で構成される。

CPU 230 は、ディスクドライブ 260 に格納されている管理プログラムをメモリ 240 に読み込んで、そのプログラムに基づくストレージ装置に対する管理処理を実行する。メモリ 240 は、例えば RAM 等で構成され、プログラムやデータ等を格納する。

ネットワークアダプタ 250 は、ホスト計算機 10 又はストレージ装置 30 と管理ネットワーク 40 を介してデータを送受信する。

ディスクドライブ 260 は、例えばハードディスク装置で構成され、データやプログラムを格納する。

ストレージ装置 30 は、コントローラ 31、ストレージキャッシュメモリ 32、共有メモリ 33、物理デバイス (PDEV) 34、電源スイッチ 35 及び電源 36 を備える。コントローラ 31 は、PDEV 34 に構成された記憶領域へのデータの格納を制御する。

ストレージキャッシュメモリ 32 は、PDEV 34 に読み書きされるデータを一時的に格納する。共有メモリ 33 は、コントローラ 31 や PDEV 34 の構成情報を格納する。PDEV 34 は、複数のディスク装置によって構成される。電源 36 は、ストレージ装置 30 の各部に電力を供給する。電源スイッチ 35 は、電源 36 からの電力の供給を ON/OFF するスイッチである。ディスク装置 (記憶デバイス) は、例えばハードディスクドライブで構成され、主としてユーザデータを格納する。記憶デバイスとしては、フラッシュメモリなどの半導体メモリからなるドライブでもよい。

コントローラ 31 は、ホストアダプタ 310、ネットワークアダプタ 320、不揮発性メモリ 330、電源制御部 340、メモリ 350、プロセッサ 360、ストレージアダプタ 370 及び共有メモリアダプタ 380 によって構成される。

ホストアダプタ 310 は、ストレージネットワーク 50 を介してホスト計算機 10 との間でデータを送受信する。ネットワークアダプタ 320 は、管理ネットワーク 40 を介してホスト計算機 10 又は管理装置 20 との間でデータを送受信する。

不揮発性メモリ 330 は、ハードディスクやフラッシュメモリで構成され、コントローラ 31 で動作するプログラムや構成情報等を格納する。電源制御部 340 は、電源 36 から供給される電力を制御する。

メモリ 350 は、例えば RAM 等で構成され、プログラムやデータ等を格納する。プロセッサ 360 は、不揮発性メモリ 330 に格納されているプログラムをメモリ 350 に読み込んで、そのプログラムに規定された処理を実行する。

ストレージアダプタ 370 は、PDEV 34 及びストレージキャッシュメモリ 32 との間でデータを送受信する。共有メモリアダプタ 380 は、共有メモリ 33 との間でデータを送受信する。

図 2 は、図 1 に示した計算機システムにおいて、ストレージ装置 30 が行う記憶領域の動的割り当ての動作を示す機能ブロック図である。PDEV 34 から RAID 構成によって RAID グループが構成される。この RAID グループから仮想デバイス (VDEV) 400 が構成される (S101)。VDEV 400 は記憶領域である複数の論理デバイス (LDEV) 500 に分割されている。PDEV 34 から構成された VDEV を、「第 1 種 VDEV」と呼ぶ。この第 1 種 VDEV に含まれる LDEV を、「第 1 種 LDEV」と呼ぶ。

ホスト計算機 10 は、ストレージ装置 30 のボリューム (論理領域) にアクセスする。ホスト計算機 10 から見えるボリュームを「ターゲットデバイス」と呼ぶ。ターゲットデバイス 700 は、第 1 種 LDEV 500 を含むボリュームにホスト計算機 10 へのパスを定義して設定される (S102)。

ストレージ装置 30 の外部に接続されている外部物理デバイス 600 を PDEV 34 と

10

20

30

40

50

同様に扱うこともできる。すなわち、RAID構成によって複数の外部物理デバイス(DEV)600から複数の第1種DEV400が構成される(S103)。第1種DEV400は一つ以上の記憶領域である第1種LDEV500に分割されている。この第1種LDEV500にホスト計算機10へのパスを設定して、ターゲットデバイス700が設定される(S104)。

また、ストレージ装置30において、第2種DEV401を設定できる。第2種DEVとは、PDEV34によって構成される第1種DEVとは異なり、アドレスと領域を持っているがPDEV34に対応する領域は持っていない仮想的なデバイスである。第2種DEVに対応するキャッシュメモリの領域の設定は可能である。この第2種DEV401には、一つ以上のLDEVを構成する。このLDEVを2種LDEV501と呼ぶ。

10

この第2種LDEV501にホスト計算機10へのパスを設定して、ターゲットデバイス701が設定される(S110)。このターゲットデバイスが仮想ボリュームに対応する。

第2種DEV及び第2種LDEVは、物理的な記憶領域がPDEVから割り当てられていない(記憶領域を持っていない)。これをホスト計算機10が使用するためには、第2種LDEVを、ストレージプール(あるいは単に“プール”という。)60に関連付ける必要がある。

ストレージプール60は、第1種LDEV500の一つまたは複数、ストレージプールの属性で纏めたグループである。第1種LDEV500は、ストレージプールに割り当てられている(S112)。第1種LDEVがプールボリュームである。

20

このストレージプールに設定された第1種LDEV500を第2種LDEV501に、アドレスを用いて、関連付けを行う(S111)。この関連付けを、“マッピング”あるいは、“割り当て”という。

これによって、ホスト計算機10が、仮想ボリュームに対してライト・リードアクセスが可能となる。また、第1種LDEVと第2種LDEVとのマッピングを変更することで、仮想ボリュームの記憶領域を変更できる。

図3は、ストレージ装置30のメモリ350ブロック図である。メモリ350には、プロセッサ360によって読み込まれて実行される各種プログラムや、LDEVの設定に関する構成情報351及びストレージプールの設定に関するプール情報352が格納される

30

。コマンド制御プログラム3501は、ホスト計算機10又は管理装置20からのコマンドを解釈し、そのコマンドに規定された処理を実行する。構成制御プログラム3503は、ストレージ装置30の構成を制御する。ディスクI/Oプログラム3505は、PDEV34へのアクセスを制御する。プール制御プログラム3507は、ストレージプールを設定する。

構成情報351は、ストレージ装置30のDEV及びLDEVに関する設定を格納する。プール情報352は、ストレージプールに関する設定を格納する。

構成情報351は、アドレス管理テーブル3511、LDEV管理情報3512、DEV管理情報3514を含む。

40

アドレス管理テーブル3511は、ターゲットデバイスとLDEVとDEVと物理デバイスとのアドレスのマッピング情報を格納する。アドレス管理テーブル3511は、ターゲットデバイス-LDEVマッピング情報35111、LDEV-DEVマッピング情報35112及びDEV-PDEVマッピング情報35113を含む。LDEV管理情報3512は、LDEVに関する情報を格納する。DEV管理情報3514は、仮想論理ボリュームに関する情報を格納する。

プール情報352は、POOL管理情報3521、POOL-VOL管理情報3522、VVOL-DIR3523、PSCB3524を含む。POOL管理情報3521は、ストレージプールの設定を格納する。POOL-VOL管理情報3522は、ストレージプールのボリュームの情報を格納する。VVOL-DIR3523は、ストレージプール

50



の L D E V のアドレスの割り当てを格納する。P S C B 情報は、ストレージプールの L D E V のアドレスの情報を格納する。

さらに、メモリ 3 5 0 は、ストレージプールに属する複数の L D E V ( プールボリューム ) 5 0 0 に P D E V から記憶領域を均等に割り当てるための割り当て処理プログラム 3 5 0 8 と、複数のプールボリューム間で記憶領域を移動させて、複数のプールボリュームの記憶容量を均等化させるためのリバランス処理プログラム 3 5 0 9 と、を有している。

割り当て処理プログラムがホスト計算機から I O がきてページを割り当てる際に、どのプールボリュームからどの頻度で割り当てるかは、仮想ボリュームごとに定義した割り当て順序テーブル (あるいは選択順序テーブル) によって定義されている。

図 4 は、V D E V 管理情報の説明図である。V D E V 管理情報は、V D E V 固有情報 3 5 1 4 1 から構成される。V D E V 固有情報 3 5 1 4 1 は、V D E V 番号 ( V D E V # ) 3 5 1 4 2、エミュレーションタイプ 3 5 1 4 3、総サイズ 3 5 1 4 4、残サイズ 3 5 1 4 5、デバイス属性 3 5 1 4 6、デバイス状態 3 5 1 4 7、設定 L D E V 数 3 5 1 4 8、L D E V 番号 3 5 1 4 9、先頭 V D E V - S L O T # 3 5 1 5 0 及び終了 V D E V - S L O T # 3 5 1 5 1 で構成される。

V D E V # 3 5 1 4 2 は、V D E V の識別子である。エミュレーションタイプ 3 5 1 4 3 は V D E V のエミュレーションタイプの識別子である。総サイズ 3 5 1 4 4 は、V D E V に設定されている総サイズである。残サイズ 3 5 1 4 5 は V D E V の未使用の領域のサイズである。

デバイス属性 3 5 1 4 6 は、V D E V に定義されている属性の識別子である。その V D E V が第 1 種 V D E V である場合には第 1 種 V D E V を示す識別子が格納され、その V D E V が第 2 種 V D E V であり、仮想ボリュームに設定されている場合には第 2 種 V D E V を示す識別子が格納される。

デバイス状態 3 5 1 4 7 は、V D E V の状態を示す識別子である。V D E V の状態は、正常、閉塞、障害閉塞等がある。閉塞は、パンク閉塞など、障害の発生意外の要因によって閉塞されていることを示す。障害閉塞は、デバイスの何れかに障害が発生しておりそのために閉塞されていることを示す。

設定 L D E V 数 3 5 1 4 8 は、V D E V に設定されている L D E V の総数である。L D E V 番号 3 5 1 4 9 は V D E V に設定されている L D E V の番号が格納される。先頭 V D E V - S L O T # 3 5 1 5 0 は、設定されている L D E V の物理的な先頭のスロット番号の識別子である。

終了 V D E V - S L O T # 3 5 1 5 1 は、設定されている L D E V の物理的な最終のスロット番号である。この L D E V 番号 3 5 1 4 9、先頭 V D E V - S L O T # 3 5 1 5 0 及び終了 V D E V - S L O T # 3 5 1 5 1 は、L D E V 数と同じ数だけ L D E V 番号毎に設定される。

図 5 は、L D E V 管理情報の説明図である。L D E V 管理情報は、V D E V 固有情報 3 5 1 2 1 から構成される。L D E V 固有情報 3 5 1 2 1 は、L D E V 番号 ( L D E V # ) 3 5 1 2 2、エミュレーションタイプ 3 5 1 2 3、サイズ 3 5 1 2 4、先頭スロット番号 3 5 1 2 5、終了スロット番号 3 5 1 2 6、パス定義情報 3 5 1 2 7、デバイス属性 3 5 1 2 8、デバイス状態 3 5 1 2 9、プログラム使用状況 3 5 1 3 0 0 及び P O O L - I D 3 5 1 3 0 1 で構成される。

L D E V # 3 5 1 2 2 は、L D E V の識別子である。エミュレーションタイプ 3 5 1 2 3 は L D E V のエミュレーションタイプの識別子である。サイズ 3 5 1 2 4 は、L D E V に設定されている総サイズである。

先頭スロット番号 3 5 1 2 5 は設定された L D E V の先頭のスロット番号の識別子である。終了スロット番号 3 5 1 2 6 は設定された L D E V の最終のスロット番号である。パス定義情報 3 5 1 2 7 は、ホスト計算機 1 0 に定義されたパスの識別子である。

デバイス属性 3 5 1 2 8 は、L D E V の属性の識別子である。L D E V が第 1 種 L D E V である場合には第 1 種 L D E V を示す識別子が格納され、L D E V が第 2 種 L D E V である場合には第 2 種 L D E V を示す識別子が格納される。また、L D E V がストレージ

10

20

30

40

50

ールに設定されている場合は、プール属性を示す識別子が格納される。

デバイス状態 35129 は、その LDEV が所属する VDEV の状態を示す識別子である。VDEV の状態は、正常、閉塞、障害閉塞等がある。閉塞は、パンク閉塞など、障害の発生意外の要因によって閉塞されていることを示す。障害閉塞は、デバイスの何れかに障害が発生しておりそのために閉塞されていることを示す。

プログラム使用状況 351300 は、LDEV が何れかのプログラムによって処理中である場合に、そのプログラムの識別子が格納される。POOL-ID 351301 は、LDEV がストレージプールに設定されている場合に、その識別子が格納される。

図 6 は、アドレス管理テーブルの説明図である。アドレス管理テーブル 3511 は、ターゲットデバイスと LDEV と VDEV と物理デバイスとのアドレスのマッピング情報を格納する。アドレス管理テーブル 3511 は、ターゲットデバイス - LDEV マッピング情報 35111、LDEV - VDEV マッピング情報 35112 及び VDEV - PDEV マッピング情報 35113 を含む。

ターゲットデバイス - LDEV マッピング情報 35111 は、ターゲットデバイスのアドレスと LDEV のアドレスとの対応が格納される。LDEV - VDEV マッピング情報 35112 は、LDEV のアドレスと VDEV のアドレスが格納される。

VDEV - PDEV マッピング情報 35113 は、VDEV のアドレスとその RAID グループ番号（又はパリティグループ）と PDEV のアドレスが格納される。

ストレージ装置 30 は、このアドレス管理テーブルを参照することによって、ターゲットデバイスのアドレスがどの LDEV のどのアドレスに対応するかを知ることができる。また、LDEV のアドレスがどの VDEV のどのアドレスに対応するかを知ることができる。また、VDEV のアドレスがどの RAID グループに属しており、どの PDEV のどのアドレスに対応するかを知ることができる。

図 7 は、POOL 管理情報の説明図である。POOL 管理情報 3521 は、POOL 固有情報 35211 から構成される。POOL 固有情報 35211 は、POOL-ID 35212、属性/用途 35213、エミュレーションタイプ 35214、容量 35215、空き容量 35216、閾値 35217、状態 35218、POOL-VOL 数 35219、POOL-VOL デバイス番号リスト 35220、POOL を利用しているデバイス数 35221 及び POOL を利用しているデバイス番号 35222 で構成される。

POOL-ID 35212 は、POOL の識別子である。属性/用途 35213 は、そのストレージプールの属性及び用途を示す識別子である。用途は、例えばスナップショットや SYS 領域など、運用形態の用途である。

エミュレーションタイプ 35214 はストレージプールのエミュレーションタイプの識別子である。容量 35215 は、ストレージプールの総容量である、空き容量 35216 はストレージプールの未使用の領域のサイズである。

閾値 35217 は、ストレージプールが許容する最大のデータ格納容量である。状態 35218 は、ストレージプールの現在の状態である。例えば、定義中、拡張中、有効等である。POOL-VOL 数 35219 は、ストレージプールとして設定されている LDEV の総数である。

POOL-VOL デバイス番号リスト 35220 はストレージプールとして設定されている LDEV 番号の一覧である。POOL を利用しているデバイス数 35221 は、ストレージプールの LDEV が関連付けられている第 2 種 LDEV の数である。POOL を利用しているデバイス番号 35222 は、ストレージプールの LDEV が関連付けられている第 2 種 LDEV 番号の一覧である。

図 8 は、VVOL-DIR 3523 及び PSCB 3524 の説明図である。VVOL-DIR 3523 は、仮想ボリュームの記憶領域である第 2 種 LDEV の構成の情報である。PSCB (POOL Slot Control Block) 3524 は、ストレージプールに設定された第 1 種 LDEV の構成の情報である。

前述のように、ストレージ装置 30 は、PDEV 34 から RAID 構成によって第 1 種 VDEV を構成する。この第 1 種 VDEV を、記憶領域である第 1 種 LDEV に分割する

10

20

30

40

50

。第1種LDEVがストレージプールに設定される。このストレージプールに設定された第1種LDEVによって構成されるボリュームをPOOL-VOL900とする。

また、ストレージ装置30は、仮想ボリューム(VVOL)を設定し、第2種VDEVを構成する。この第2種VDEVを記憶領域である第2種LDEVに分割する。

ストレージ装置30は、VVOL800の第2種LDEVを、POOL-VOL900の第1種LDEVに割り当てる。これによって、ホスト計算機10がアクセスする仮想ボリュームの記憶領域が、物理デバイスであるPDEV34から構成された第1種LDEVに設定される。

VVOLの構成は、VVOL-DIR3523に格納される。VVOL-DIR3523は、LDEV番号(LDEV#)35231及びエントリ35232によって構成される。

10

LDEV番号(LDEV#)35231は第2種LDEVの識別子である。エントリ35232は、第2種LDEVの構成情報である。このエントリ35232は、第2種LDEVアドレス35233及びPSCBポインタ35234から構成される。

第2種LDEVアドレス35233は、VVOLの第2種LDEVのアドレスが格納される。

PSCBポインタ35234は、第2種LDEVがPOOL-VOL900の第1種LDEVに割り当てられている場合に、その第1種LDEVの領域のポインタが格納される。なお、初期状態では第2種LDEVは第1種LDEVに割り当てられていないので、PSCBポインタ35234には「NULL」が格納される。

20

PSCB(POOL Slot Control Block)3524は、ストレージプールに設定されている第1種LDEVの情報である。このPSCB3524は、ストレージプールに設定されている第1種LDEVの-slot毎に設定される。

PSCB3524は、LDEV番号(LDEV#)35241、POOL-VOLアドレス35242、PSCB前方ポインタ35243及びPSCB後方ポインタ35244から構成される。

LDEV番号(LDEV#)35241は、POOL-VOLでの第1種LDEVの識別子である。POOL-VOLアドレス35242は、POOL-VOL900における第1種LDEVのアドレスである。

PSCB前方ポインタ35243及びPSCB後方ポインタ35244は、POOL-VOL900内の第1種LDEVの前後の-slotの識別子である。

30

また、POOL-VOL900の領域のうち、未使用の領域は、その先頭がフリーPSCBキュー35240で示される。フリーPSCBキュー35240は、次の-slotを示すPSCB3524へのポインタを含む。

ストレージ装置30は、フリーPSCBキュー35240に示されたポインタを参照して、次のPSCB3524を得る。さらに、次のPSCB3524のPSCB後方ポインタ35245を参照して、段階的にPSCB3524を辿る。そして、その未使用の領域の最終の-slotに対応するPSCB3524を得る。この最後のPSCB3524のPSCB後方ポインタ35244はフリーPSCBキュー35240である。ストレージ装置30は、フリーPSCBキュー35240を辿り、PSCB3524のポインタによって連結された集合によって、ストレージプールのPOOL-VOL900の未使用の領域を知ることができる。

40

ストレージ装置30は、ストレージプールに設定された第1種LDEVに対応するPSCB3524を設定する。具体的には、ストレージプールに設定された第1種LDEVの各-slotに対応するPSCB3524を設定し、さらにフリーPSCBキュー35240を設定する。初期状態ではストレージプールは全て未使用であるため、フリーPSCBキュー35240によって連結される集合は、ストレージプールに設定された第1種LDEVの全ての領域に対応する。

そして、ストレージ装置30は、このストレージプールの領域を使用する場合に、必要な-slot分のPSCB3524を第2種LDEVであるVVOL-DIR3523に割

50

り当てることで、当該領域が使用可能となる。複数のスロットの集号がページに相当する。ページは複数のP S C Bから特定される。ホスト装置からV V O L 8 0 0へのアクセス、V V O L 8 0 0のアクセス領域に対するP O O L - V O Lからの記憶領域の割り当てはページ単位で実行される。

具体的には、ストレージ装置30は、フリーP S C Bキュー35240を参照する。そして、第2種L D E Vに割り当てる必要な領域分(ページ)のP S C B3524を取得する。この取得したP S C B3524を、それぞれV V O L - D I R3523のエントリに割り当てる。すなわち、V V O L - D I R3523の各エントリのP S C Bポインタ35234に、対応するP S C B3524を示すポインタを格納する。

なお、割り当て済みのP S C B3524は、フリーP S C Bキュー35240の連結から外す。

10

これによって、第2種L D E Vの各ページ(スロット)が、V V O L - D I R3523の各エントリのP S C Bポインタ35234で示されるP S C B3424に割り当てられる。P S C B3524は第1種L D E Vのスロットに対応しているので、結果として、第2種L D E Vが第1種L D E Vに割り当てられ、ホスト計算機10のアクセス対象である仮想ボリュームが物理デバイスとして使用可能となる。

図9は、計算機システムの第2の形態を示すハードウェアブロック図である。本実施形態の計算機システムも、一つ以上のホスト計算機10、管理サーバ20、第1のストレージ装置125、及び、第2のストレージ装置161を備えている。第1のストレージ装置125は、第1のネットワーク121を介してホスト計算機10に接続される。第2のストレージ装置161は、第2のネットワーク123を介して第1のストレージシステム125に接続されている。一つ以上のホスト計算機10、管理サーバ20、第1のストレージ装置125及び第2のストレージ装置161は、第3のネットワーク108を介して相互に接続される。

20

第1のネットワーク121、第2のネットワーク123及び第3のネットワーク108は、いかなる種類のネットワークであってもよい。例えば、第1のネットワーク121及び第2のネットワーク123は、いわゆるS A Nであってもよい。第3のネットワーク108は、いわゆるL A Nであってもよい。

第1のストレージ装置125は、コントローラと記憶デバイス群34とを備える。コントローラは、例えば、複数のフロントエンドインタフェース127と、複数のバックエンドインタフェース137と、第1の内部ネットワーク156と、一以上のキャッシュメモリ32と、一以上の制御メモリ350と、一以上の制御プロセッサ360とを備える。

30

フロントエンドインタフェース127は、ストレージ装置125にネットワークを介して接続されたホスト計算機10又は第2のストレージ装置161と通信するためのインタフェース回路である。

従って、ストレージ装置125は、少なくとも二つのフロントエンドインタフェース127を備え、それらのうち一つは第1のネットワーク121に接続され、別の一つは第2のネットワーク123に接続される。

フロントエンドインタフェース127は、例えば、第1のネットワーク121又は第2のネットワーク123に接続されるポート129と、メモリ131と、ローカルルータ(以下、「L R」と略記する)133とを備える。L R133に、ポート129及びメモリ131が接続される。

40

L R133は、ポート129を介して受けたデータを任意の制御プロセッサ143で処理するために振り分ける。具体的には、例えば、制御プロセッサ143が、或るアドレスを指定するI/Oコマンドをその制御プロセッサ143に実行させるようにL R133を設定する。その設定に従って、L R133が、I/Oコマンド及びデータを振り分ける。

バックエンドインタフェース137は、P D E V34と通信するためのインタフェース回路である。バックエンドインタフェース137は、例えば、P D E V34に接続されるディスクインタフェース141と、メモリ135と、L R139と、を備える。L R139に、ディスクインタフェース141及びメモリ135が接続される。

50

第1の内部ネットワーク156は、例えば、スイッチ（例えばクロスバスイッチ）或いはバスで構成される。第1の内部ネットワーク156に、複数のフロントエンドインタフェース127、複数のバックエンドインタフェース137、一以上のキャッシュメモリ32、一以上の制御メモリ350及び一以上の制御プロセッサ143が接続されている。これらの要素間の通信は、第1の内部ネットワーク156を介して行われる。

コントローラの構成要素であるフロントエンドインタフェース127、バックエンドインタフェース137、キャッシュメモリ32、制御メモリ350及び制御プロセッサ360に、第2の内部ネットワーク（例えばLAN）155が接続され、その第2の内部ネットワーク155に、保守管理端末153が接続される。

保守管理端末153は、第3のネットワーク108にも接続されており、ストレージシステム125を保守又は管理する計算機である。ストレージシステム125の保守員は、例えば、保守管理端末153（又は、その保守管理端末153と通信可能な管理サーバ111）を操作して、制御メモリ350に記憶される種々の情報を定義することができる。

第2のストレージ装置161は、コントローラ165及びPDEV163を備える。コントローラ165は、例えば、ネットワークアダプタ162と、ホストアダプタ164と、キャッシュメモリ172と、制御メモリ171と、プロセッサ167と、ストレージアダプタ169と、を備える。

ネットワークアダプタ162は、第3のネットワーク108に接続され、管理サーバ111と通信するインタフェースである。

ホストアダプタ164は、第2のネットワーク123に接続され、第1のストレージ装置125と通信するインタフェースである。ホストアダプタ164は、例えば、第1のストレージ装置125のフロントエンドインタフェース127と同様ののものであってもよい。

制御メモリ171は、種々のコンピュータプログラム及び情報を記憶するメモリである。キャッシュメモリ172は、第1のストレージ装置125からのI/Oコマンドに従って読み出される又は書き込まれるデータを一時記憶するメモリである。

プロセッサ167は、制御メモリ171に記憶されている種々のコンピュータプログラムを実行する。少なくとも、プロセッサ167は、第1のストレージ装置125からのI/Oコマンドに従って、キャッシュメモリ172及びPDEV163に対するデータの書き込み及び読み出しを制御する。

PDEV163は、物理記憶デバイスであり、例えば、第1のストレージ装置のPDEV34と同様ののものであってもよい。あるいは、PDEVは、テープ記憶媒体であってもよい。

本実施形態の第1のストレージ装置125は、いわゆる外部接続機能を備える。第2のストレージ装置161は、この機能によって第1のストレージ装置125に外部接続されている。ここで、外部接続について説明する。

既に説明したように、第1のストレージ装置125は、一つ又は複数の論理ボリュームをホスト計算機10に提供する。各論理ボリュームは、ホスト計算機10によって一つの記憶デバイスと認識される。例えば、第1のストレージ装置125が提供する論理ボリュームが、第1のストレージ装置125内のPDEV34に対応付けられてもよい。その場合、第1のストレージ装置125は、論理ボリュームへのライトコマンドを受信すると、その論理ボリュームに対応付けられたPDEV34にデータを格納する。このような論理ボリュームは、以下の説明において通常ボリュームとも記載される。

あるいは、第1のストレージ装置125が提供する論理ボリュームは、第2のストレージ装置161内のPDEV163に対応付けられてもよい。この場合、第1のストレージ装置125は、論理ボリュームへのライトコマンドを受信すると、その論理ボリュームに対応付けられたPDEV163にデータを書き込むためのライトコマンドを生成し、生成したライトコマンドを第2のストレージ装置161に送信する。第2のストレージ装置161は、第1のストレージ装置125から受信したライトコマンドに従って、データをPDEV163に格納する。このような論理ボリュームは、以下の説明において仮想ボリューム

10

20

30

40

50

ームとも記載される。

このように、第1のストレージ装置125が提供する論理ボリュームに格納されるデータを、実際には第1のストレージ装置125の外部に接続された第2のストレージ装置161に格納する機能が、外部接続機能と呼ばれる。

コントローラ125は、ストレージ制御処理を確立するクラスタ1251を複数備える。各クラスタは内部ネットワーク156を有し、複数のクラスタ内部ネットワーク156がクラスタ間のネットワーク1561によって接続される。

したがって、一つのクラスタの制御プロセッサ360は他のクラスタにアクセスすること、例えば、他のクラスタのキャッシュメモリ32のデータをリード・ライトすることができる。複数のクラスタ間のネットワーク1561は、パス、スイッチによって構成される。

10

図10Aは、ターゲットデバイスとLDEV及びPDEVとの説明図である。ホスト計算機10は、ストレージ装置30に設定されたターゲットデバイス700をアクセスする。ターゲットデバイス700の記憶領域は、第1種LDEV500である。

第1種LDEV500は、物理ボリュームからRAID構成によって構成された第1種VDEV400の領域である。第1種VDEV400が一つ以上の第1種LDEV500に分割されている。

また、ホスト計算機10は、ストレージ装置30に設定された仮想ボリュームをターゲットデバイス701としてアクセスする。このターゲットデバイス701の仮想的な記憶領域は第2種LDEV501である。第2種LDEV501は、ストレージプールとして設定されている第1種LDEV500に関連付けられている。すなわち、第1種LDEVと第2種LDEVとの関連付けは、アドレスのマッピングによって管理されている。

20

ホスト計算機10が、仮想ボリュームであるターゲットデバイス701にアクセスする。この仮想ボリュームの記憶領域は第2種LDEV501である。ストレージ装置30は、第2種LDEV501へのアクセスを受け付けると、第2種LDEV501に関連付けられている第1種LDEV500をアクセス先に変換する。ホスト計算機からのライトデータは、直接プールボリュームのVDEVに格納される。

図10Bは、図9に示した第2の計算機システムの図10Aに対応するブロック図である。第1種VDEVと第2種VDEV同士はアドレスをマッピングすることによって関係付けられており、ホストからのライトデータは、PDEVに直接格納される。

30

図11に、1つの仮想ボリューム(第2種LDEV)と複数のプールボリューム(第1種LDEV)との間で、均等に、アクセス対象及び記憶領域の単位としてのページ(1, 2, 3)が割り当てられている様子を示す。仮想ボリュームにはプールが割り当てられており、プールには複数のボリュームがある。仮想ボリュームのページは、複数のボリュームのページが対応している。

ホスト装置から仮想ボリュームにアクセスがあると、仮想ボリュームのページには、プールボリュームのページが割り当てられる。

図11は、仮想ボリュームに、ホスト装置から(1)、(2)、(3)の順番でアクセスがあると、均等割当て処理プログラム3508(図3)は、ページ(1)をプールボリューム#0に、ページ(2)をプールボリューム#1に、ページ(3)をプールボリューム#2に順番に割り当てる処理を行う。

40

図11の状態から、プールにボリューム#3を追加して、ページ割り当てをプールボリューム間で均等にしない場合、仮想ボリューム中の未割り当て領域へのライトに対して、空き領域を多く有するプールボリューム#3からの仮想ボリュームに対する割り当てが多くなる。当然、空き領域を多く有するRAIDグループへのアクセスが集中するため、ストレージ装置のホスト装置からのアクセス処理性能が低下する。

そこで、リバランス処理プログラム3509(図3)は、プールにボリュームを追加すると、追加したプールボリュームと既存のプールボリュームとの間で記憶容量の不均衡が生じるとして、追加したプールボリュームを使用する前、各プールボリューム間で記憶容量が均等になるように、既存のプールボリュームから追加したプールボリュームにページ

50

を移動することとした。複数のプールボリューム間でページを移動することを、ページの再配置、あるいはページのリバランスということにする。図 1 2 は、リバランスの過程を示すブロック図である。

なお、プール管理プログラムがプール情報を管理しており、プールボリュームが追加されると、プール管理プログラムがこの追加を検出する。そして、その契機で、リバランス処理プログラムが起動する。

リバランス処理プログラム 3 5 0 9 は、まず、プールに追加されたプールボリューム # 3 において割り当てに使用されていない未使用のページを確保する。

そして、割り当て済みページ数が他のプールボリューム # 1、# 2 より大きいプールボリューム # 0 のページ ( 4 ) のデータを、プールボリューム # 3 の未使用のページにコピーする ( S 1 2 0 2 )。

10

リバランス処理プログラムは、プールボリューム # 0 に割り当てられていたページ ( 4 ) の割り当てを解放して、これをフリーキューに接続する。さらに、仮想ボリュームのページ ( 4 ) の領域の割り当てを、プールボリューム # 0 のページからプールボリューム # 3 のページ ( 4 ) に変更する ( S 1 2 0 0 )。ページ間のデータのコピーをマイグレーションという。この過程が、ページの移動であり、ページの再配置である。ページの移動をページマイグレーションともいうことにする。

これらの処理の結果、複数のプールボリューム全体に対するページの割り当てが均等になるため、複数のプールボリュームそれぞれに対する I / O のレートが同程度になり、ストレージシステムのホストからのアクセスに対する処理性能がプールボリュームの追加後も良好に維持される。

20

図 1 2 では、複数のプールボリューム間で割り当てられたページ数が、複数のプールボリュームの容量それぞれが同一であることを前提に、同数になるようにした。

複数のプールボリューム間で割り当てページを均等にする上で、プールボリュームの容量を勘案して、複数のプールボリューム間で、全ページ数 ( プールボリュームにある、割り当て可能な全ページ数 ) 中割り当てに使用されたページの割合 ( ページ使用率 ) を均等にすることが重要である。ページ数の代わりに記憶容量で計算しても同じことである。

図 1 3 はそのことを示すブロック図である。符号 1 3 0 0 は、各ボリュームに割り当てられているページの全容量であり、1 3 0 2 は、ボリューム間でマイグレーションされるページの容量である。プールボリュームの容量比は、プールボリューム # 0 : # 1 : # 2 が 1 : 3 : 2 である。したがって、ページの使用率が同じ割合になるように、プールボリューム # 1 および # 3 からプールボリューム # 2 にそれぞれページが再配置される。

30

図 1 6 A はプールを管理するテーブルであり、図 1 6 B はプールボリュームを管理するテーブルである。これらのテーブルは、プール情報 3 5 2 ( 図 3 ) の一部として、メモリ 3 5 0 に記録されている。これらのテーブルは管理情報から作成される。ストレージシステムのリバランス処理プログラム 3 5 0 9 は、これらテーブルを参照して、図 1 2 及び図 1 3 で説明するリバランス処理系を構築する。

図 1 6 C は、フリーキュー単位を管理するテーブルである。仮想ボリュームのアクセス領域に対する未使用ページ ( プールボリューム ) の割り当ては、既述のように、フリーキューによって管理される。フリーキューはプールボリューム毎に設定されている。

40

フリーキューにはプールボリュームの複数の未使用ページが F I F O 構造で接続されている。未使用ページが仮想ボリュームへ割り当てられる際は、未使用ページが順番にフリーキューから解放されて、仮想ボリュームからのマッピング、データコピー先として処理される。

このテーブルは、フリーキューの識別番号 ( # ) と、フリーキューが空き領域を管理する対象となるプールボリューム # と、フリーキューに連結するフリー領域の個数と、フリーキューが所属するプール # からなる。このテーブルも、プール情報の一環としてメモリ 3 5 0 に記録されている。

図 1 7 は、ホスト装置から仮想ボリュームへのライトアクセスに対して、プールに属するプールボリュームから仮想ボリュームにページを割り当てる順序を示した説明図である

50

。ここでの説明は、記憶容量の比が2：3：2である複数のプールボリュームを持ったプールを前提としている。

仮想ボリュームのページをプールボリュームに均等に割り当てるためには、プールボリュームに追加されたボリュームを含めて、ボリュームに対してページの割り当て順序を定めることが必要である。この割り当て順序は、管理情報としてメモリに登録されている（図18参照）。

割り当て処理プログラム3508は、この管理情報を参照して、複数あるフリーキューの中から目的のフリーキューのアドレスを算出して、フリーキューに繋がる未使用ページのアドレス（ポインタ情報：PSCB）を取得し、ここに仮想アドレスのページに対して送られた、ホスト装置からのライトデータを格納する。

これを可能にするために、フリーキューのアドレスとプールボリュームの識別情報とは1：1に対応している。例えば、フリーキュー#0はプールボリューム#0のフリー領域を管理しており、フリーキュー#1はプールボリューム#1のフリー領域を管理しており、フリーキュー#2はプールボリューム#2のフリー領域を管理している。

図17は、プールボリューム#0に連続して2回ページを割り当て、次いで、プールボリューム#1に連続して3回ページを割り当て、さらに、プールボリューム#2に連続して2回ページを割り当て、以後これを繰り返すことを示している。

図18は、この割り当て順序を管理するテーブルであり、図17に示す割り当ての方式が登録されている。このテーブルもプール情報としてメモリに登録されている。割り当て処理プログラム3508は、この管理テーブルを参照して、プールボリュームに対するページの割り当ての順次制御処理を管理する。

このプログラムは、ソフトウェアカウンタを備え、どのプールボリュームまでページを割り当てたか否かを管理する。カウンタはページの割り当ての都度インクリメントされることから、カウンタの位置によって、現在のページ割当て対象となっているプールボリュームが分かる。図18は、カウンタがプールボリューム#1（フリーキュー#1）のところにあることを示している。

ストレージ装置がページの再配置（図12、図13）を実行することの契機となる事象は、図12及び図13に示すように、プールボリュームをプールに追加することのほか、図14に示すように、プールからプールボリュームの削除（Shrink）、複数のプールボリューム間で使用されているページの全体のページ数に対するバランス、あるいはプールボリュームの容量に対する使用量のバランス、その他、GUIを介してユーザ側からの要求によるものがある。

図14において、リバランス処理プログラム3509は、プールの空き容量を一定時間監視する（S33001，S33002）。プール全体の使用量（割り当てに使用された全ページの容量）が閾値を越えたか否かを判定する（S33003）。

閾値を越えると、管理装置20又は保守端末は、プールに新たにプールボリュームを追加する処理を行う。ストレージ装置は、図5乃至図7に示す管理情報をプールボリュームの追加に合わせて更新する（S33004）。その後、リバランス処理に移行する（S33020）。リバランス処理を、後に、図15に基づいて説明する。

一方、プールボリュームの削除も、リバランス処理の契機となる。GUIを介して管理ユーザからリバランス指示があると、ストレージシステムはリバランス処理を実行する。GUIは管理サーバ20あるいは保守端末に存在する。

さらにまた、リバランスプログラムの監視機能は、所定タイミングごとに、プールの各プールボリュームについて、使用されているページの全ページ数に対するバランスをチェックして（S33009）、これを閾値と比較する（S3310）。閾値より大きい場合には、バランスが崩れているとして、ストレージ装置はリバランス処理を行う。

図15はリバランスを示すフローチャートである。リバランス処理プログラムが、リバランスの対象となったプールについてリバランスの可否を判定する（S7007）。この判定は、例えば、図3のプール情報352を判定して行われる。プールの状態（図7の35218）が閉塞状態にあるときには、リバランスができないと判定される。

10

20

30

40

50



リバランスが可能と判定されると、リバランス処理プログラムは前処理を行う（S7010）。たとえば、ページマイグレーション（ページの割り当ての変更あるいは移動）移動元のプールボリュームと移動先のプールボリュームとを決定する、ページマイグレーションの処理が適用されるプールボリュームの選択順序（割り当て順序）を決定する、そして、ページマイグレーション後のプールボリュームの容量をシミュレートするなどである。

プールボリュームがプールに追加されると、プール管理情報によってプールが管理され、ストレージ装置は、プールボリュームの容量比の見直しを行う。

次いで、リバランス処理プログラムは、プール内でマイグレーションすべきページ数を算出する（S7011）。その後、ページのマイグレーション処理を行う（S7012）。

10

このとき、後述するが、マイグレーションの際に、複数のプールボリュームに存在する固定パターン（例えば、ページの全てのデータ“0”のもの）を破棄して、プールボリュームの容量を縮小するための処理を同時に行う。

ページマイグレーションの際、リバランス処理プログラムは、リバランス処理の中断条件の発生の有無を判定する。例えば、リバランスの対象となったページにホスト計算機からI/Oがあった場合には、リバランスを中断する。リバランスの中断の要因がなくなった後リバランス処理を再開してもよい。

次いで、リバランス処理プログラムは、ページマイグレーションの進捗を管理する（S7013）。この管理によって、リバランス処理プログラムは、リバランスの進捗状況を管理ユーザに対して見せることができる。

20

図19は、図15のマイグレーションするページ数の算出（S7011）の詳細を示すフローチャートである。ここでのページ数の算出の方式は、複数のプールボリューム間でバランスされたページ数を理想のページとし、これと現在使用されているページ数の差を埋めるようにすることである。

リバランス処理プログラムは、HDP VOL（仮想ボリューム）に割り当てられている各プールボリュームでのページの使用量を検索する。

図19の説明において、次のことを前提とする。

HDP VOL # i のプール使用量全体（プールに属する全てのボリュームにおいて割り当てられている（使用されている）ページの総数）：A [ i ]

HDP VOL # i のプールボリューム # j 使用量全体（使用するページの総数）：B [ i ] [ j ]

30

プールボリュームの # j の容量（総ページ数）：C [ j ]

プール全体の容量（総ページ数）：D

追加プールVOL全体の容量（総ページ数）：E

まず、リバランス処理プログラムは、プールボリューム追加前に、HDP VOL [ i ] に割り当てられているプールVOL [ j ] が使用してページ数を算出し、これをB [ i ] [ j ] に格納する（S1900）。B [ i ] [ j ] はメモリ350のワーク領域に設定されている。

次いで、プールボリュームをプールに追加後に、HDP VOL [ i ] がプールボリューム [ j ] を使用すべきページ数を次のようにして求める（S1902）。

40

HDP VOL # i のプール使用量全体 × プールボリューム [ j ] の理想の使用率 = A [ i ] × C [ j ] / ( D + E )

プールボリュームの理想の使用率とは、プールボリュームに格納可能な総ページ数に対する、プールボリュームに割り当てられたページ数の割合であって、複数のプールボリューム間で均等にページが割り当てられていることを前提とした値である。

プールボリュームボリューム [ j ] について移動が関係するページ数は、プールボリューム追加前と追加後の差であって、次のとおりである（S1904）。

ページ数 = B [ i ] [ j ] - A [ i ] × C [ j ] / ( D + E )

上記計算値がプラス（+）の値の場合は、プールボリューム [ j ] は移動元ボリュームとなり、計算値がマイナス（-）の値の場合は、プールボリューム [ j ] は移動先ボリ

50

ームとなる。

リバランス処理プログラムは、移動ページ数を算出したら、プール番号毎に、移動ページ数、プールボリュームが移動元か移動先かを専用のテーブルに登録する。

図20は、図19の処理によって算出されたページ数とページの移動の様子を示すブロック図である。ページの移動は、一つのストレージプールグループに属する複数のプールボリュームと、これに対応する複数の仮想ボリュームと、において、各仮想ボリュームについてページの移動を算出した結果に基づく。プールボリュームのシンボルの中の数値は、プールボリュームにおいて割り当てに使用されているページ数を示している。

図20に示す様に、仮想ボリューム#0について、複数のプールボリューム間でページの移動はない。仮想ボリューム#1、#2について、図20に示す矢印の方向で、矢印の上

10

に示してある数のページがマイグレーションされる。  
リバランス処理プログラムは、プール単位でページの移動を行う。すなわち、図20に示す全てのページ移動を行う。なお、仮想ボリューム単位でページの移動を行ってもよい。たとえば、仮想ボリューム#1について、複数のプールボリューム間でページの移動を行い、仮想ボリューム#2については、ページの移動を実行しないようにしてもよい。

図21は、図20に示すページの移動形態を実現するための管理テーブルである。この管理テーブルはメモリ350の所定領域に格納される。管理テーブルは仮想ボリューム毎に存在し、図19のステップS1906の処理に基づいて作成される。

仮想ボリューム毎で各プールボリュームが移動元となるかあるいは移動先となるかが登録されており、移動元のプールボリュームから移動先のプールボリュームについて移動されるページ数が移動元及び移動先のプールボリュームにそれぞれ登録されている。

20

リバランス処理プログラムは、S7011(図15)の処理でこのテーブルを作成・登録し、S7012でこのテーブルを参照してマイグレーションを実行する。

図22は、バランス見直し(図14のS33009)を実現するためのページ割り当てバランス管理テーブルである。割り当て管理プログラム3508がページをプールボリュームに割り当てていく過程で、リバランス処理プログラム3509は、この管理テーブルに必要な事項を登録する。このテーブルはメモリの所定領域に格納される。この管理テーブルは仮想ボリューム毎に存在する。

このテーブルにおいて、理想の割り当てページ数とは、プールボリュームについて現在割り当てられているページ数に対して理想となる割り当てページ数であり、複数のプールボリュームの容量比を勘案して決定される。プールボリューム#0:#1:#2:#3(容量比)は、2:3:2:2である。仮想ボリュームに割り当てられているページの総数が18であるので、容量比を勘案すると、理想の割り当てページ数の比は、4:6:4:4になる。

30

そして、割り当てバランスは、理想の割り当てページ数に対する実際の割り当てページ数の比である。割り当てバランスが1を超えるプールボリュームは、過剰にページが割り当てられていることを示し、逆に、これが1未満であるプールボリュームは、過少にページが割り当てられていることを示している。

均等割り当て処理プログラム、又は管理装置は、割り当てページ管理テーブルを定期的に参照して、割り当てバランスの良否を判定する。例えば、1.0から±0.35の範囲が割り当てバランスが良いとすると、プールボリューム#0のバランスが悪く、他のプールボリュームのバランスはよい、ということになる。

40

バランスが悪いと判断されたプールボリューム数が全体プールボリュームに対して\*\*%(閾値:図14のS33010)を超えたら、プール全体のバランスが悪いと判断し、リバランス処理プログラムがページについてマイグレーション処理を開始する。割り当てページ管理テーブルのチェックは、ストレージ装置の保守端末の管理プログラムが実行してもよい。

一方、割り当てバランスの良否の判断は、プールボリューム個々について行われてもよい。たとえば、割り当て比が、1.0から±0.85の範囲外のプールボリュームが少なくとも一つあったら、ストレージシステムは、プール全体のバランスが悪いと判断し、リ

50

バランスを開始する。

図23は、図15のマイグレーション処理(S7012)の詳細を示すフローチャートである。この処理は、複数の仮想ボリュームが存在する場合は、各仮想ボリュームについて行われる。

リバランス処理プログラムは、仮想ボリュームの中から、関連ページを検出する。移動対象となるのはプールボリュームのページであるが、アドレス管理テーブル3511を参照することにより、移動対象となるページに対応する関連ページを仮想ボリュームの全ページの中から特定することができる。

リバランス処理プログラムは、仮想ボリューム中の先頭から全ページのそれぞれについて、関連ページであるか否かの判定を行う。どのページまでチェックしたかを管理するためにソフトウェアカウンタを用いる。ページのチェックが進む都度、カウンタをインクリメントする。そこで、リバランス処理プログラムは、カウンタの初期化を行う(S2300)。

10

次いで、リバランス処理プログラムは、仮想ボリューム中のページを先頭ページから選択する(S2302)。次いで、選択したページにプールボリュームの移動対象ページが割り当てられているか否かを判定する(S2304)。

リバランス処理プログラムは、この判定を肯定すると、アドレス管理テーブル3511に基づいて移動対象ページが属するプールボリュームの識別情報を検出する(S2306)。

次いで、図21のテーブルを参照して、プールボリュームが移動元か移動先かを判定する(S2308)。リバランス処理プログラムは、プールボリュームが移動先である場合には、移動元プールボリュームからページ移動するための空き領域を移動先プールボリュームに確保する(S2310)。

20

ページが移動されるプールボリュームとしては、割り当てバランスが低いものが選択されるが、複数のプールボリュームが移動先として競合した場合には、追加されたプールボリューム、次いで、移動先となる上での優先度が高いプールボリュームから順に選択される。移動先優先度テーブルには、各プールボリュームについて移動先となる優先度が登録されている。

リバランス処理プログラムは、移動先プールボリューム、移動元プールボリューム、移動対象ページ数を決めると、ページの移動のスケジュールを決定し、これをメモリ350のマイグレーション実行ジョブに登録する(S2312)。

30

このスケジュールは、例えば複数のページの移動を伴う場合、どのような順番でページの移動を行うかを定めるためのものである。このスケジュールは、S2302にて順次選択されたページがある都度、新たな移動先ボリュームなどが決定される都度更新登録される。

次いで、リバランス処理プログラムは、カウンタのカウント値をインクリメントし(S2316)、S2302にて選択していないページが未だ存在するか否かを判定する(S2318)。

未検討のページが存在する場合にはステップS2302にリターンし、さらに、ページが存在しない場合にはS2312において順次更新されたスケジュールにしたがって、マイグレーションが実行される(S2320)。なお、マイグレーション処理(S2320)をS2312の後に実行するようにしてもよい。

40

また、S2320の処理がS2312とS2316の2316の間に入ってもよい。

図24は、固定データパターンの破棄を伴う、マイグレーション処理の詳細を示すフローチャートである。このフローチャートは移動元ページが複数ある場合には各移動元ページについて適用される。

ストレージアダプタ370は、リバランス処理プログラムにしたがって、移動元プールボリュームから移動元ページの全データをスロット単位で、キャッシュメモリ32にステージングする(S2400)。

50

図 2 6 にこの様子が示されている。プールボリューム # 1 は移動元であり、プールボリューム # 2 は移動先である。S 2 6 0 0 は、移動元ページがキャッシュメモリの第 1 の領域 3 2 A にステージングされる様子を示している。符号 2 6 0 1 はスロットである。ステージングはスロット単位で実行される。符号 2 6 0 1 A は移動元ページであり、符号 2 6 0 1 B は移動先ページである。

ストレージアダプタ 3 7 0 内の D R R 回路 ( D a t a R e c o v e r a n d R e c o n s t r u c t 回路 ) は、スロット単位で移動元ページ 2 6 0 1 A のデータをチェックし、データが “ 0 ” か否かをチェックする ( 図 2 4 : S 2 4 0 2 - S 2 4 0 4、図 2 6 : S 2 6 0 2 )。

この判定は、移動元ページのデータが固定パターンの一つ ( 全データが “ 0 ” ) であるか否かを判定するためのものである。データが固定パターンの場合にはページのマイグレーションは行わず仮想ボリュームの移動対象ページに対するマッピングを、固定パターンデータを格納している専用領域に変更する。

10

D R R が “ 1 ” の存在を検出すると、ステップ 2 4 1 4 に移行して、キャッシュメモリ内メモリコントローラは、移動元ページデータを、第 1 領域 3 2 A から第 2 領域 3 2 B へ、キャッシュメモリ間コピーを行う ( 図 2 4 : S 2 4 1 4、図 2 6 : S 2 6 0 4 )。そして、ストレージアダプタ 3 7 0 は、移動元キャッシュメモリ 3 2 A をクリアする ( 図 2 4 : S 2 4 1 6 ) する。

次いで、第 2 領域 3 2 B のデータを移動先プールボリューム # 2 の移動先ページ 2 6 0 1 B にデステージ ( 図 2 6 : S 2 6 0 6 )、移動元ページ 2 6 0 1 A を未使用ページに変更し、さらに、仮想ボリュームのページに対する割り当てを移動元ページから移動先ページに変更するなどのマイグレーション処理を行う ( 図 2 4 : S 2 4 1 8 )。

20

D R R が判定しているスロットのデータが “ 0 ” の場合は、ストレージアダプタ 3 7 0 は、スロット単位の全データのチェックが終了したか否かを判定し ( S 2 4 0 6 )、これを否定判定する場合には、ステップ S 2 4 0 2 に戻り、同一スロットの他のデータについて、“ 0 ” か否かをチェックする。

スロットの全部のデータについてチェックが終わると、ステップ S 2 4 1 0 に移行し、移動元ページに残りのスロットがあるか否かをチェックして、これを肯定判定する場合には、次のスロットについて既述のチェックを行う。

ページに属する全てのスロットについて判定が終わると、ステップ S 2 4 1 2 に移行して、全データが “ 0 ” からなる移動元ページは、マイグレーションが必要でないとして、未使用ページに変更する。

30

ストレージ装置が図 2 4 に示す処理を実行することにより、ページを再配置する際、固定パターンデータを持つページを使用済みの管理から解放して、固定パターンデータが重複して仮想ボリュームに割り当てられることを解消し、ひいては、ホスト装置は、プールボリュームの記憶領域を効率よく利用することができるようになる。

既述の実施形態では、0 データの存在のチェックを、ストレージアダプタ 3 7 0 内の回路 ( D R R ) によりパリティ生成時についでに処理するとしたが、専用ソフトウェアで行う、又は、専用回路で行うものであってもよい。

図 2 4 で説明するマイグレーションの過程で、移動元のページに対して、ホスト計算機から I / O またはストレージ装置内コピー I / O が発行されると、ストレージ装置は、I / O 処理の方を優先させてページマイグレーションを中止する。

40

すなわち、図 2 5 に示す様に、仮想ボリュームにリード又はライト I / O が供給されると、プロセッサは、移動元ページに対する論理アドレスが含まれているか否かをチェックする ( S 2 5 0 0 )。

移動元ページに対する論理アドレスが含まれていることは、仮想ボリュームと移動元ページとの間のパスが切断されてなく、ページマイグレーション処理は終了していないことを意味するため、プロセッサは、移動元ページのデータが格納されている、キャッシュメモリの第 2 の領域 ( 図 2 6 の 3 2 B ) のデータをクリアし、さらに移動先ページに対してクリア処理を行う ( S 2 5 0 4 )。

50

ここでのクリア処理とは、仮想ボリュームにマッピングされている場合には、マッピング管理情報を削除することであり、また、移動先ページが使用ページとして管理されているのであれば、その管理情報を削除することである。

次いで、プロセッサは、S 2 5 0 6において、図 2 4のマイグレーション処理をスキップする(終了する)。なお、リードI/Oの場合は、この処理が終了するまで、マイグレーション処理を一時停止し、その後再開するようにしてもよい。

図 2 6で説明したように、キャッシュメモリ 3 2には、第 1領域(移動元用のキャッシュ領域) 3 2 Aと、第 2領域(移動先用のキャッシュ領域) 3 2 Bとがあり、両者の差分が無くなるまで、第 2領域から移動元ページのデータが移動先ページ 2 6 0 1 Bにコピーされないため、移動元ページにI/Oが入ったタイミングでページの移動処理を中止することができる。

10

これに対して、図 3 1に示す、キャッシュメモリの構成(図 9の形態)では、ローカルルータ(LR) 1 3 3のDRRが0データの検出を行い、プロセッサは、キャッシュメモリ 3 2の一つのキャッシュ領域 3 2 Aにステージングされたデータを他のキャッシュ領域を利用することなく直接移動先ページ 2 6 0 1 Bにデステージするため、ページのマイグレーションは、図 2 6のキャッシュメモリの構造のものより高速に進むことができる半面、移動元ページ 2 6 0 1 Aにホスト装置からI/Oのアクセスがあっても、ページのマイグレーションを中止できづらくなる。

図 3 1の場合でも、0チェックしたデータをキャッシュメモリにためこみ、結果0でなければプールボリューム# 2に格納する。図 3 1の斜め書きの方が高速である。

20

次に、図 2 4のS 2 4 1 8のページマイグレーション処理をフリーキューに関連して説明する。この説明のために、図 3 2に示すマイグレーションを前提にする。すなわち、プールボリューム# 0のページ# 1 0 0 0のデータが、プールボリューム# 3のページ# 2 0 0 0にコピーされ、仮想ボリューム(HDP Vol # 0)の前者のページに対するマッピングを后者のページのマッピングに変更される過程を、図 3 3を用いて説明する。

リバランス処理プログラム 3 5 0 9は、移動中キューを確保し、これをフリーキュー# 3に接続し、次いで、プールボリューム# 3に割り当てられているフリーキュー# 3に接続するフリー領域の中から、移動先ページとなる対象領域 3 3 0 2を選択し、このポインタ情報(Pool - Vol # 3、# 2 0 0 0) 3 3 0 3を移動中キュー(キュー# 0) 3 3 0 0の最後に接続する(S 3 3 0 0)。移動先ページとなる領域を、フリーキューとは別キューで管理するのは、移動先ページとなる領域が仮想ボリュームに割り当てられないようにするためである。

30

コピージョブからページデータのマイグレーションが正常に終了したとの通知をリバランス処理プログラムが受けると、V VOL DIRにあるページ(プールボリューム# 0、ページ# 1 0 0 0)へのマッピング 3 3 0 6の情報である、ポインタ情報 3 3 0 4をクリアして、これを、移動先ページのアドレス(Pool - Vol # 3、# 2 0 0 0)へのポインタ情報に変更し(S 3 3 0 2)、移動元ページのアドレスをフリーキュー#のキュー構造の適切な位置に挿入する。さらに、移動先ページを移動中キューから解放する。

図 3 4に示す様に、移動元ページから移動先ページへのデータマイグレーションの過程が、ホスト装置からのI/Oなどによって、これが中断(3 4 0 0)されると、図 3 5に示す様に、リバランス処理プログラムは、移動中キューを確保し、移動先キューに対して設定されたポインタ情報 3 3 0 4をクリアし、変更前の情報に戻す。

40

さらに、移動先となるページ 3 3 0 3がもともと存在していたプールボリュームのフリーキュー# 3を確保し、このフリーキューに対して移動先ページ 3 3 0 3がフリー領域として挿入される位置を検索する。

リバランス処理プログラムは、移動中キュー 3 3 0 0をフリーキュー(Queue # 3)に接続し、移動先ページ 3 3 0 3を移動中キュー 3 3 0 0から解放し、さらに、移動先ページをフリーキュー(Queue # 3)に接続する。

以上のように、ストレージ装置はフリーキューとは別キューである、移動中キューを利用することにより、ページの再配置を可能にするとともに、ページ再配置に競合するホス

50

トI/Oなどの要因がストレージ装置に加わっても、I/Oの処理を優先させることができる。

次に、図27A、図27B、図27Cを利用して、仮想ボリュームとプールボリュームについて、固定データパターンに対する割り当て処理の詳細をここで説明しておく。固定データパターンとして、ページ内の全データが“0”のパターンを例とする。

図27Aは、仮想ボリューム(DP VOL)の作成を示すブロック図である。ストレージ装置は、仮想ボリュームの未割り当て領域(未使用ページ)2700の全てを、プールボリュームの0データの専用領域2702に割り当てる。

0データ専用領域は、プール内の少なくとも一つのボリュームに設けてもよいし、キャッシュメモリに設けてもよい。ホスト計算機から仮想ボリュームの未割り当て領域にリード要求がきたら、ストレージ装置は、プールボリュームの0データ専用領域をリードして0データをホスト計算機に返す。

10

図27Bに示す様に、ストレージ装置が、仮想ボリュームの未割り当て領域にホスト装置からライト要求を受けると、プールボリュームの空きページを選択して、ここにライトデータを格納する。

ストレージ装置は、ライト要求2708を受けた仮想ボリュームのページ2706の割当先を「0データ専用領域」2702から割当て領域2704に変更する。すなわち、ストレージ装置は、割り当て2710を解放して、新たな割り当て2712を管理テーブルに設定する。ライト要求が全“0”データの場合には、プールボリュームの割り当て済み領域2704に全“0”データを上書きする。

20

次いで、図27Cに示す様に、ストレージ装置は、仮想ボリュームの先頭ページから順番に、割り当てられているプールボリュームのページのデータが全て0であるかどうか検索し(2720)、全データが“0”のページ2722を検出したら、仮想ボリュームのページ2724の割り当て先を現在割当てられているページ2722から「0データ専用領域」としての専用ページ2702に変更する。

全データが“0”である、プールボリュームのページを使用済み領域ではなく空き領域として管理する。この“0データページ破棄処理”は、図24に示す様に、ページのリバランス処理に付帯して実行されたが、ページリバランス処理とは独立して実行されてもよい。

この時、“0データページ破棄処理”に伴って、複数のプールボリューム間でページ割り当て比に不均衡が生じる場合には、図14のS3309、及び、図22で説明したように、リバランス処理が適用される。

30

図28は、0データ専用領域へのリード処理を示すフローチャートである。プロセッサ360は、仮想ボリュームへのホスト装置からのリード要求が、未割り当て領域に対するものであるか否かを判定する(S35001)。これを否定判定すると、プールボリュームの割り当て領域のアドレスを検索し、リード要求に対して、検索されたページのデータを応答する(S35005)。

一方、リード対象が未割り当て領域の場合には、全0データをホスト装置に対して応答するが、ステップS35002において、仮想ボリュームに割り当てられているプールボリュームの数をチェックする。

40

この判定は、0データ専用領域の設定方法が複数あることに起因している。1つ目は、0データ専用領域をプールの先頭プールボリュームに1つだけに設ける。2つ目は、プール内の全てのプールボリューム0データ専用領域を設ける。

0データ専用領域を設定したプールボリューム数が少ない(1つ目の方法、2つ目の方法でもプールボリューム数が少ないなど)と、PDEV(記憶デバイス)へのリード負荷が大きくなるため、ストレージ装置が0データを作成してホスト計算機に応答する(S35005)。ストレージ装置は0データを作成して、例えば、メモリ(キャッシュメモリ)上で作り出す。0データ専用領域を設定したプールボリューム数の大小は、規定値との比較による。

一方、0データ専用領域が設定されたプールボリュームが多い場合には、リード負荷が

50

複数のプールボリュームに分散されるため、ストレージ装置はプールボリュームの0データ専用領域に対応する物理領域をリードして0データをホスト装置に返答する(S35003)。

なお、記憶デバイスが高速なリード性能を持つもの(例SSD)であれば、リード負荷を無視することができるので、プールボリュームの数の如何に拘わらず、プロセッサは、記憶デバイスに常時0データを読みに行くようにすればよい。プール毎に、記憶デバイスのリード特性を管理することにより、これは可能である。

また、プールに、高回転速度なHDDと低回転速度なHDDとが混在して割り当てられている場合、高回転速度であるHDDに設定されたプール又はプールボリュームに0データ専用領域を設定し、これに仮想ボリュームの未割当て領域をマッピングすればよい。

図29は、仮想ボリュームの未割当て領域にホスト装置からライト要求が発行された場合の処理(図27B)について説明するものである。ストレージ装置は、仮想ボリュームのライト対象が未割当て領域であるか否かを判定する(S2900)。これを否定判定すると、仮想ボリュームのライト対象ページに割り当てられているページ(プールボリュームの)のアドレスを検索し、ページにライトデータを上書きする(S2902)。

一方、前記判定を肯定すると、プールボリュームの空き領域からページを選択し(S2904)、ここにライトデータを格納する(S2906)。ライト要求を受けた仮想ボリュームのページは、“0データ専用領域”からプールボリュームのページに向けてマッピング管理情報を変更する(S2908)。

既述の実施形態では、移動元プールボリュームに割り当てられているページを移動先プールボリュームにマイグレーションする際に、移動対象ページが固定データパターンである場合には、移動対象ページをマイグレーションすることなく解放し、仮想ボリュームのページを移動先プールボリュームの固定データパターンを格納している領域にマッピングした。

さらに、ストレージ装置は、プールにプールボリュームが追加された際に、複数のプールボリューム間で、ページをマイグレーションする契機で、“0”データ専用領域に関するリバランスを行うようにしてもよい。

すなわち、図30Aに示す様に、追加プールボリューム3000に0データ専用領域3002を設定し、既存のプールボリューム3008の0データ専用領域3006に対して設定されていた、仮想ボリューム3004の未割当て領域3005からのマッピング3001が、追加プールボリューム3000の専用領域3002のマッピング3003に変更される。

したがって、仮想ボリュームの未割当て領域から0データ専用領域へのマッピングが、追加したプールボリュームを含めた複数のプールボリュームに対して均等に行われる。

また、図30Bに示す様に、プールボリューム3100が削除されると、削除されたプールボリュームの使用済みページが残りのプールボリュームに均等に再配置されるが、この際、削除されたプールボリュームの専用領域3102に対する仮想ボリュームの未割当て領域3104、3106からの3108も、残りのプールボリュームの専用領域3102、3114に均等に変更される(3110、3112)。

このように、仮想ボリュームのページに対する0データ専用領域の割り当てを複数のプールボリュームに分散させることによって、仮想ボリュームにライトアクセスがあった際での、仮想ボリュームの未使用領域の割り当てを0データ専用領域から解放し、次いで、未使用領域をプールボリュームの未使用領域に割り当てる処理を、複数のプールボリュームにわたって分散させることができる。

図36は、マイグレーション進捗管理(図15のS7013)のための管理テーブルである。リバランス処理プログラムはこの管理テーブルへ管理情報を登録し、登録した管理情報を更新する。マイグレーションの進捗状況は、画面への表示によって管理ユーザに供給される。管理情報には、プール毎に「マイグレーションを行っているかないか」の情報があり、具体的には次のものがある。

リバランス中：リバランス処理実行中。

10

20

30

40

50

リバランス未：(1)リバランスしていない、(2)リバランス対象外。

リバランス未完了：(1)リバランス処理が中断(障害、ユーザ指示など)され、計画の途中までしか実行していない、(2)リバランス処理の計画は最後まで実行されたが、スキップされたページがあった、(3)電源OFFした、など。

さらに、リバランス処理を失敗した場合は、リバランス中断したことを画面に示し、ユーザが再度リバランス処理を実行できるタイミングを与えるようにすること、マイグレーション対象のプールについて、HDP VOL総数のうち、何個のHDP VOLをマイグレーション完了したかの情報を提供するようにしてもよい。

次に、図14のプールボリュームを削除する処理(Shrink“シュリンク”)を詳細に説明する。図37は、リバランスを考慮したプールボリュームのシュリンク処理を説明するフローチャートである。プールボリュームのシュリンクは、GUIを介してユーザ指示による。

10

図38は、GUIによってユーザに提供される情報の例であり、プール及びプールボリュームの特性が含まれている。ユーザはこの情報に基づいて、シュリンクするプールボリューム、このプールボリュームが属するプール、シュリンクする容量、プールボリュームがシュリンクされた後での、プールにおけるページの使用率などを指定する。

ページの使用率とはプール内のプールボリュームに割り当て可能な全ページ数に対する仮想ボリュームに実際に割り当てられているページ数の比である。シュリンクはプールボリュームに対して実行される。

ユーザからのシュリンク条件がリバランスプログラムに指示されると(S3700)、シュリンクの対象となるプールがシュリンクできるものであるか、あるいはそうでないかを判定する(S3702)。

20

プール管理情報、プールボリューム管理情報にアクセスして、これらの状態を参照する。例えば、プールやプールボリュームが閉塞されていたり、あるいはプールに属するプールボリュームにホスト装置からのI/Oやストレージシステム内のコピーI/Oが存在する場合は、管理サーバあるいはストレージシステムの保守端末のGUIがユーザにシュリンク不可を通知する(S3720)。

シュリンクが可能であることを判定すると、リバランス処理プログラムは、ステップS3704において、ユーザからプールボリュームの指定があったか否かをチェックする。プールボリュームが指定された場合、指定されたプールボリュームがプールから削除されても、プール全体におけるページの使用量あるいは使用率が閾値を超えないと、ステップS3706に移行してシュリンクした後でのプール・プールボリュームの特性を予想して、これをGUIによってユーザに表示する(S3708)。

30

図39はシュリンクの予想結果の出力例である。シュリンク前のプールの容量、使用量、空き容量と割合は、プール管理情報からデータを採取する。空き容量の割合は、プールの容量と空き容量から算出する。

シュリンク後のプールの容量、使用量、空き容量と割合は、ユーザが指定したシュリンクプールボリューム、又はストレージ側でシュリンク対象と決定したプールボリューム全ての容量、使用量、空き容量をプールボリューム管理テーブルから算出し、プールの情報から減算する。

40

S3704で、削除すべきプールボリュームがユーザから指定されない、と判定すると、ストレージ装置がユーザから削除する容量を受け取った際、どのプールボリュームを削除するかを決める。

例えば、期待する使用量が入力された場合、現状の使用量から期待する使用量になるために、削除する容量を算出し、ストレージ側でどのプールボリュームを削除するかを決める。ストレージ側でプールボリュームを選択する場合、割り当てページ数の少ないプールボリュームから選択すれば、移動量を少なく抑えることができる。

対象プールボリュームをシュリンクしてもプール使用量が閾値(たとえば、全体容量の80%以上)を超えないかを、ストレージ装置がチェックする。

使用量が閾値を超えることになる削除の要求は却下される。プールボリューム#\*\*を

50



ユーザが入力した場合、ユーザ指定のプールボリュームを削除することにより、使用量が閾値を超えることになる場合は、ユーザの削除要求を却下する。

また、シュリンクを実施することで、プールボリューム数が所定値（例：8）以下になると、リバランス機能が十分に発揮されないため、残存するプールボリューム数が所定値異化になる場合は、ユーザへその旨の警告を出す。

図39の値を画面に出力し、ユーザは入力した条件、即ち、指定したプールボリュームをシュリンクしてもよいかを確認し、これを許容する場合、「続行」をユーザがGUIに入力する（S3710）。これによりシュリンク計画が決定され、ストレージ装置は、シュリンク処理とリバランス処理を開始する。

ステップS3702において、プールボリュームが指定されない場合は、ストレージ装置がシュリンクするプールボリュームを決定する（S3722）。図40は、それを実現するための処理の詳細を示したフローチャートである。

リバランス処理プログラムは、シュリンクする容量を超えない範囲でシュリンクするプールボリュームを決定する。シュリンクするプールから、割り当てられたページの容量が小さい順にプールボリュームを並べる（S4000）。

次に、割り当てられたページの容量が最も少ないプールボリュームの容量とシュリンクすべき容量（当初の容量又はシュリンク後ではシュリンクの残量）とを比較し、プールボリュームの容量がシュリンクすべき容量を超える場合には、ステップS4008に移行して他のプールボリュームを選択して、再度ステップS4002を行う。

一方、シュリンクすべき容量よりプールボリュームの容量が少ない場合には、このプールボリュームをシュリンク対象ボリュームとして選択（S4004）し、シュリンクすべき容量から選択したプールボリュームの容量を減じ、シュリンクの残量用を更新し（S4006）、ステップ4008に進み、次にページ使用量が小さいプールボリュームについて、ステップ4002の処理を行う。図40に示す処理はシュリンクの対象となったプールに属する全てのプールボリュームについて行われる。

図37のステップS3710にしたがい、ユーザがシュリンクを許容する入力をストレージ装置に行うと、以後、プールボリュームの削除処理を伴った、ページマイグレーション処理の前処理が行われる。

次いで、ステップS3714に移行し、シュリンクを伴って行うマイグレーションが適用されるページ数を算出する。さらに、ステップS3716において、ページマイグレーションを行い、S3718において、マイグレーションの進捗管理が行われる。S3712 - S3718は、図15において説明した処理とほぼ同様である。

図41及び図42は、プールボリュームの削減を伴ったリバランスを説明するためのブロック図である。図41は、プールからプールボリューム#j+1が削除されることを示している。残存するプールボリュームの容量比に応じて、プールボリューム#j+1が有している使用済みページ数（4ページ）を、図41に記載するような割合で既存のプールボリュームに移動させる。

図41に示す様に、第1には、削除するプールボリュームから複数のページが、移動先となるプールボリュームの容量比に応じてプールボリュームに分配される。

この際、ストレージ装置、図42に示すように、複数のプールボリュームについて現実の割り当てページ数が、理想のページ数になるように、プールボリュームのシュリンクによるページの移動4202と、リバランス処理によるページの移動とをバランスさせて行う。

図43は、既述の実施形態の変形例を示すものであり、図9に示すストレージ装置が複数のクラスタを備えていることを示すハードウェアブロック図である。第1のクラスタ1251が、第1の仮想ボリューム（HDP VOL#0）に対するリバランスを制御し、第2のクラスタが第2の仮想ボリューム（HDP VOL#1）に対するリバランスを制御する。

プール#0、プール#1、プール#3は、それぞれのクラスタ内、あるいは複数のクラスタを跨いで形成される。リバランスは、これら複数のプール全体に対して同時に行って

10

20

30

40

50

もよいし、個々のプールについて個別に行うようにしてもよい。

例えば、プールVOLグループ#2で使用しているときに、プールVOLグループ#0にプールVOLが追加された場合、プールVOLグループ#2内のプールVOLグループ#1に含まれるプールVOLのデータをプールVOLグループ#0内のプールVOLへ移動し、プールVOLグループ#2をプールVOLグループ#0にする。もしくは、プールVOLグループ#1内のプールVOL数を減らす。

すなわち、図43の実施形態は、プールはクラスタ間にまたがって作成される。そのプールに所属するプールボリュームのうち、仮想ボリュームと同じクラスタにあるプールボリュームにページが極力割り当てられるようになっている。

つまり、実体と同じクラスタにあれば、クラスタ間をまたがったデータ転送が発生しないためである。それでも、プールボリュームが足りなくなったら、別のクラスタのプールボリュームが使用されるが、極力クラスタからはみでる範囲を減らそうとする。

プールVOLがクラスタにまたがっている状態でプールVOLが追加されたら、クラスタからはみでて格納しているデータをクラスタ内に入るようにリバランスする。

図44はさらに他の実施形態を示すものであり、仮想ボリューム(HDP VOL)の属性に、ホスト装置からのアクセスに対する応答の高速性を示す「性能」を追加した例を示すブロック図である。例えば、「性能」には、「超高速」「高速」「低速」のような識別子を登録する。プールボリュームの属性にも、この「性能」を設定することができる。

管理ユーザは、仮想ボリュームの属性を考慮して、この属性に合うようにプールボリュームを仮想ボリュームに対して設定する。例えば、「超高速」の仮想ボリューム#1には、SSDから作成したプールボリュームを割付ける。「高速」の仮想ボリューム#2には、SASやFCから作成したプールボリュームを割付ける。

「低速」の仮想ボリューム#3には、SATAから作成したプールボリュームを割付ける。その他、外部ストレージのプールボリュームを割り付ける仮想ボリュームがあってもよい。ストレージ装置は、SSDのプールボリュームの範囲でリバランスを行う。SSDのプールボリュームが追加された場合は、SSDのプールVOLを使用する仮想ボリュームがリバランスする。

これは、図43と同様の考え方である。図43ではプールにまたがったかどうかでグルーピングしているが、図44はHDDの種類でグルーピングしている。1つのプールの中にSSDのグループ(プール#1とする)やSASやFCのグループ(プール#2とする)など存在する。プールにSSDのプールボリューム(=HDD)が追加されたら、プールの中のプール#1内でリバランスし、プール#2などの他のところには影響しない動きとなる。

既述の実施形態において、プールボリュームがプールから削除されると、このプールボリュームはページ割当て順序情報管理テーブルから削除されるために、このプールボリュームの仮想ボリュームのアドレスにリンクする新たなページが割当てられることはないが、プールからプールボリュームが完全に削除されるまでは、ホスト装置からのリードアクセスがこのプールボリュームに供給される。

ストレージ装置の制御メモリにリバランスのオン/オフフラグ設定情報を、管理ユーザは管理装置又は保守装置を介して設定可能である。この設定情報をオンにしておくと、契機が発生しても、ストレージ装置はページのリバランス処理を行わない。この設定情報は、プール毎、仮想ボリューム毎に設定することもできる。

既述の実施形態では、ページリバランス契機の判定、リバランス実施の可否をストレージ装置が行ったが、これを管理装置が行うようにしてもよい。前者の場合は、ストレージ装置が本発明の計算機システムになり、後者の場合は、管理装置とストレージ装置との組み合わせが計算機システムになる。

10

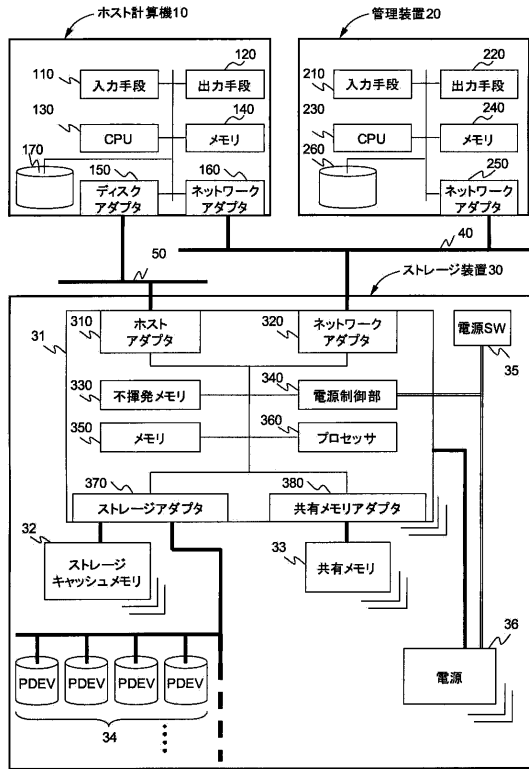
20

30

40

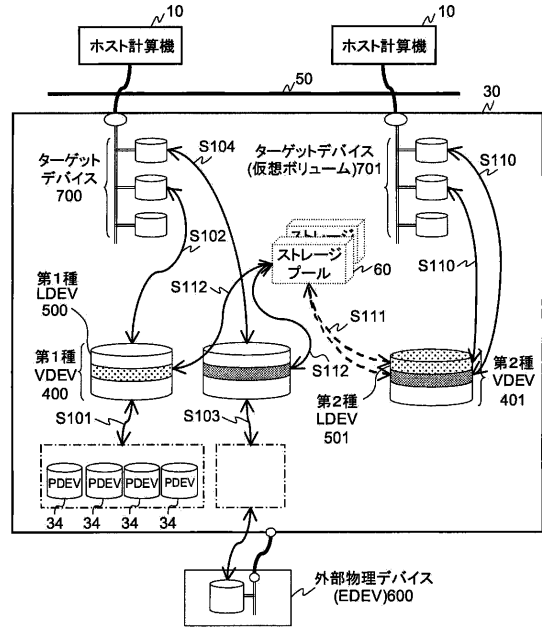
【図1】

図1



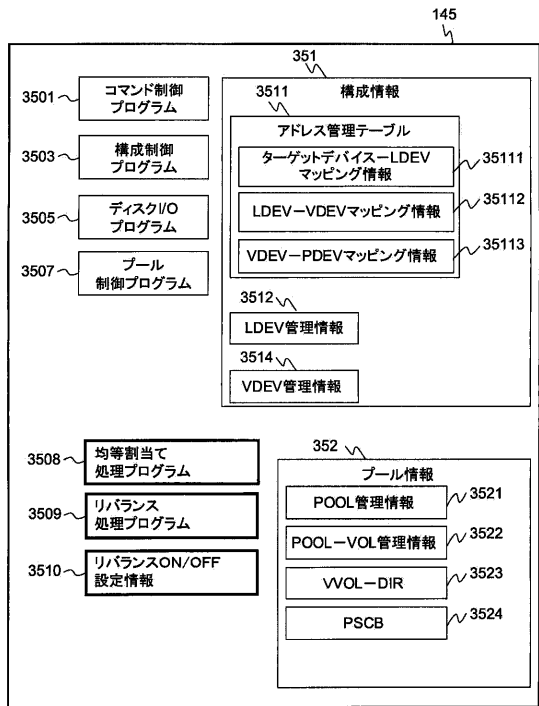
【図2】

図2



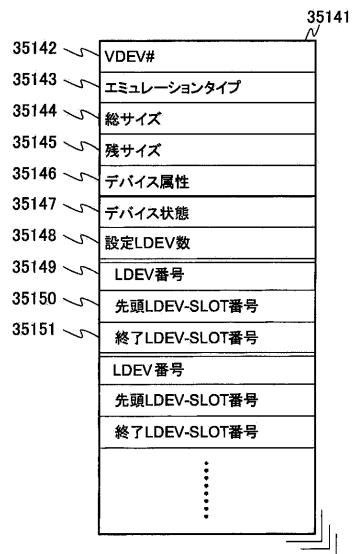
【図3】

図3



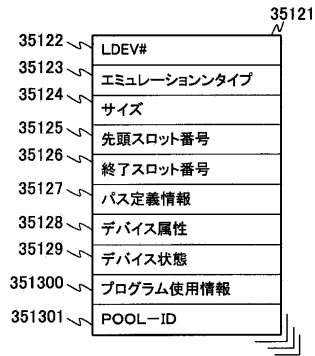
【図4】

図4



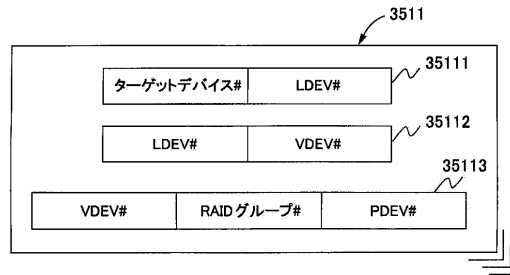
【 図 5 】

図5



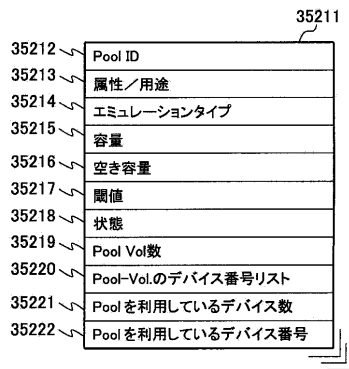
【 図 6 】

図6



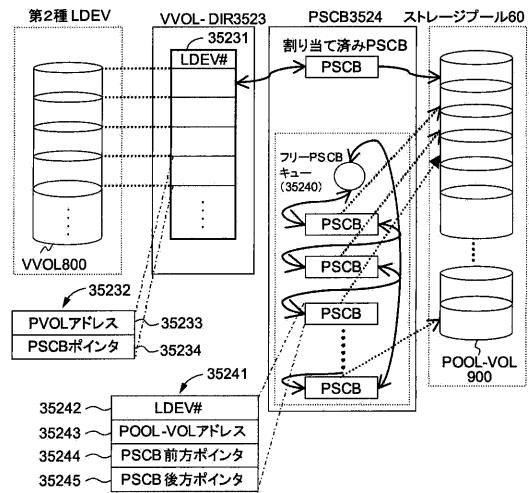
【 図 7 】

図7



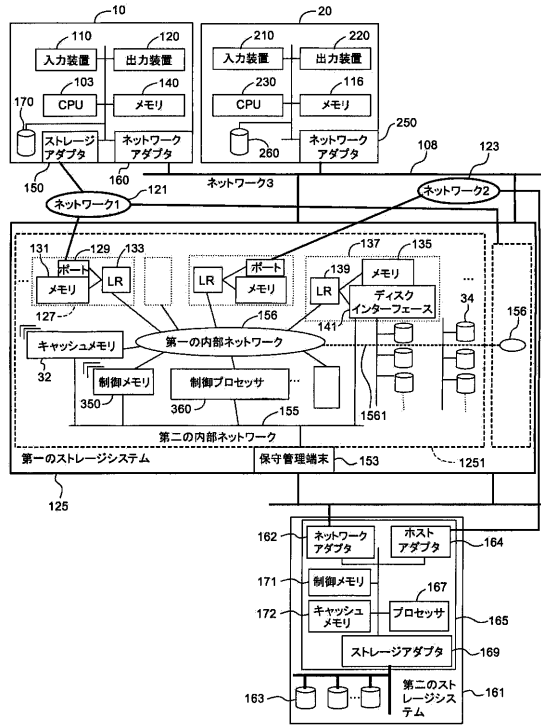
【 図 8 】

図8



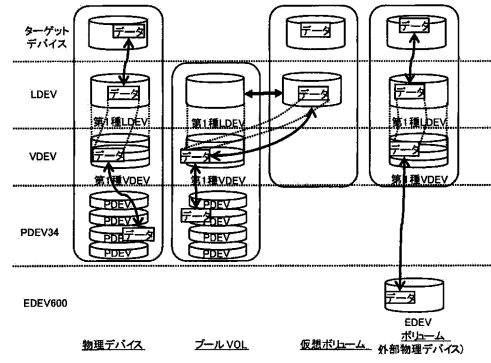
【図9】

図9



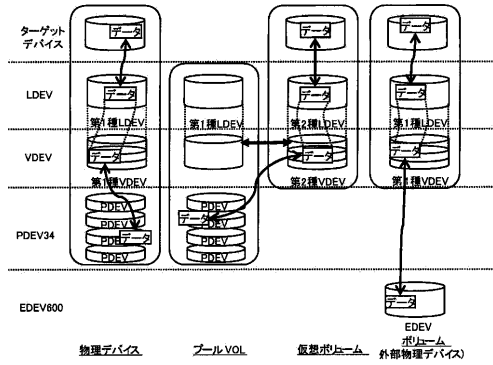
【図10A】

図10A



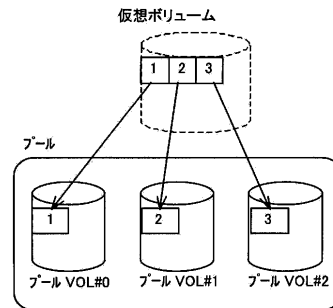
【図10B】

図10B



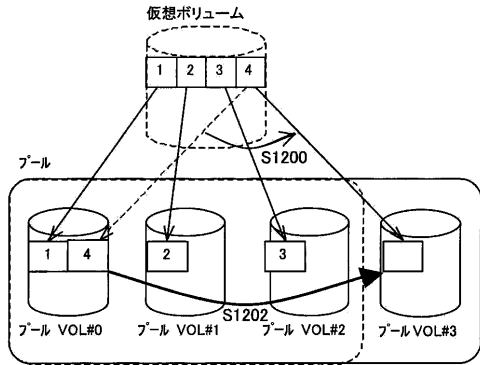
【図11】

図11



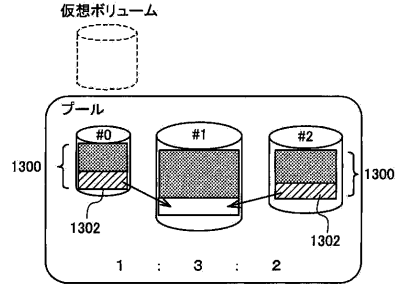
【図 1 2】

図12



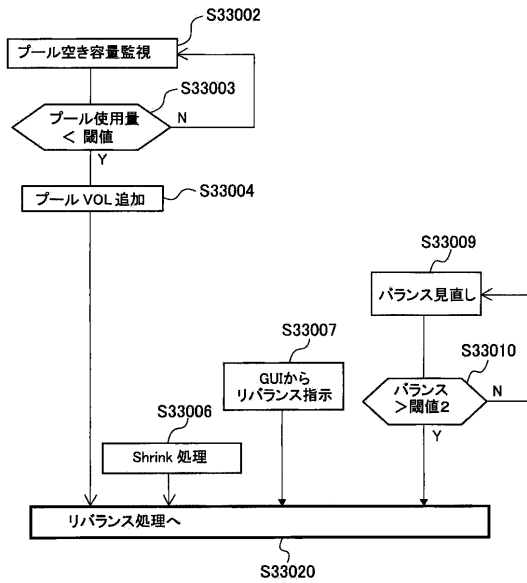
【図 1 3】

図13



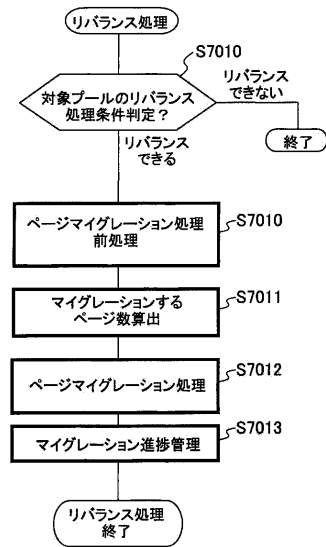
【図 1 4】

図14



【図 1 5】

図15



【図16A】

図16A

プール管理テーブル

| プール# | 容量<br>ページ数 | 使用量<br>ページ数 | 所属する<br>プールVOL# |
|------|------------|-------------|-----------------|
|------|------------|-------------|-----------------|

【図16B】

図16B

プールVOL管理テーブル

| プール# | 容量<br>ページ数 | 使用量<br>ページ数 | 所属する<br>プールVOL# |
|------|------------|-------------|-----------------|
|------|------------|-------------|-----------------|

【図16C】

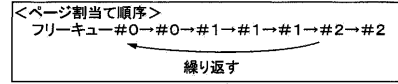
図16C

フリーキュー単位管理テーブル

| フリーキュー# | フリーキューが<br>管理する<br>プールVOL# | フリーキューに<br>つながる個数 | 所属する<br>プール# |
|---------|----------------------------|-------------------|--------------|
|---------|----------------------------|-------------------|--------------|

【図17】

図17



【図18】

図18

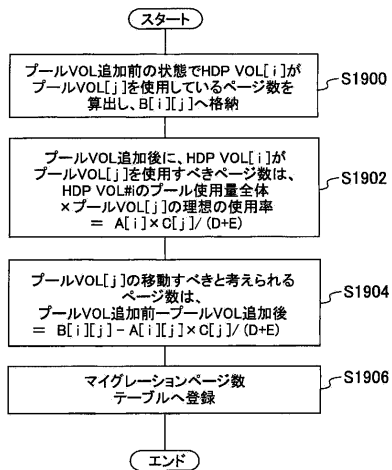
割当て順序管理テーブル

| フリーキュー# | 割当比 | 割当回数 |
|---------|-----|------|
| 0       | 2   | 2    |
| 1       | 3   | 2    |
| 2       | 2   | 0    |

← カウンタ

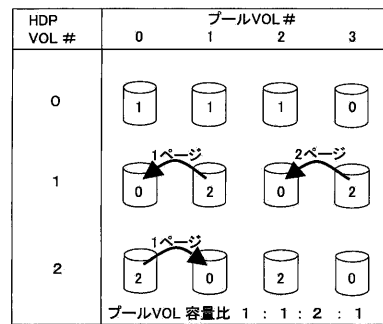
【図19】

図19



【図20】

図20



【図 2 1】

図21

| HDP # 0 のプール VOL |        |        |
|------------------|--------|--------|
| プールVOL #         | 移動ページ数 | 移動元か先か |
| 0                | 0      | —      |
| 1                | 0      | —      |
| 2                | 0      | —      |
| 3                | 0      | —      |

| HDP # 1 のプール VOL 移動量 |        |        |
|----------------------|--------|--------|
| プールVOL #             | 移動ページ数 | 移動元か先か |
| 0                    | 1      | 移動先    |
| 1                    | 1      | 移動元    |
| 2                    | 2      | 移動先    |
| 3                    | 2      | 移動元    |

| HDP # 2 のプール VOL 移動量 |        |        |
|----------------------|--------|--------|
| プールVOL #             | 移動ページ数 | 移動元か先か |
| 0                    | 1      | 移動元    |
| 1                    | 1      | 移動先    |
| 2                    | 0      | —      |
| 3                    | 0      | —      |

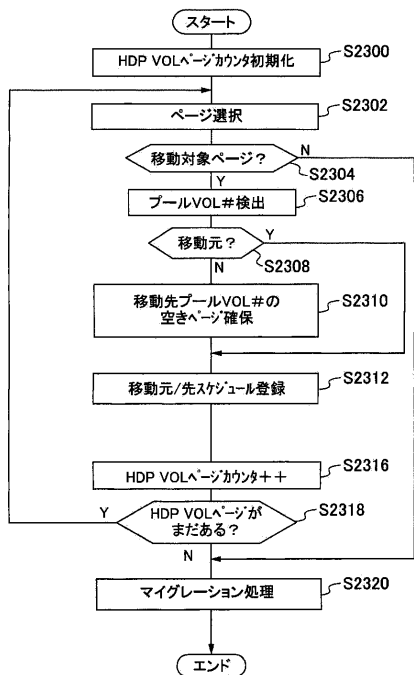
【図 2 2】

図22

| HDPVOL #0に割り当たったページのプール |     |        |            | VOL の内訳 |
|-------------------------|-----|--------|------------|---------|
| プールVOL #                | 割当比 | 割当ページ数 | 理想の割当てページ数 | 割当てバランス |
| 0                       | 2   | 2      | 4          | 2/4     |
| 1                       | 3   | 6      | 6          | 6/6     |
| 2                       | 2   | 5      | 4          | 5/4     |
| 3                       | 2   | 6      | 4          | 6/4     |
| プールからのページ割当て総数          |     | 18     |            |         |

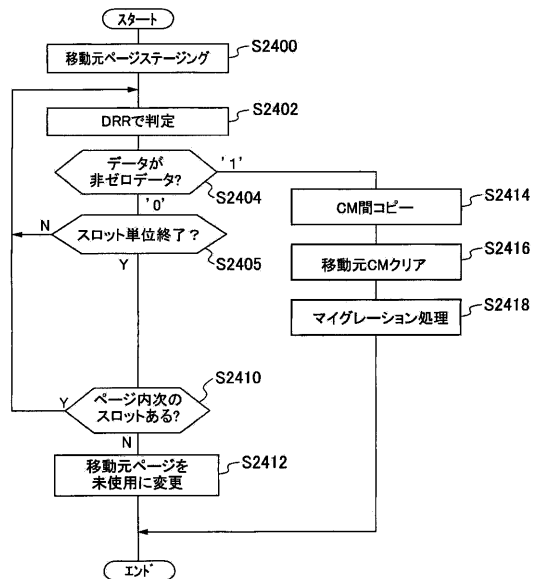
【図 2 3】

図23



【図 2 4】

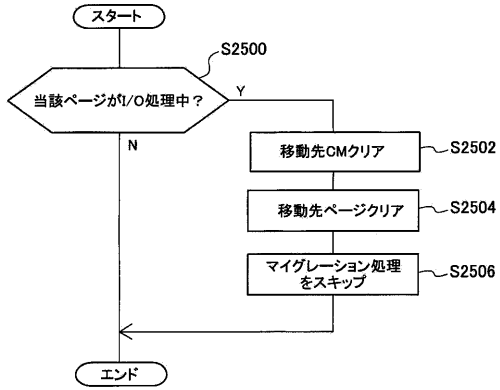
図24





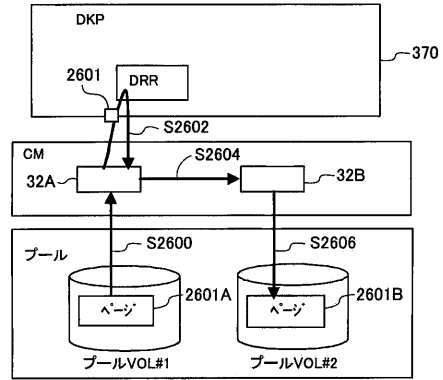
【図25】

図25



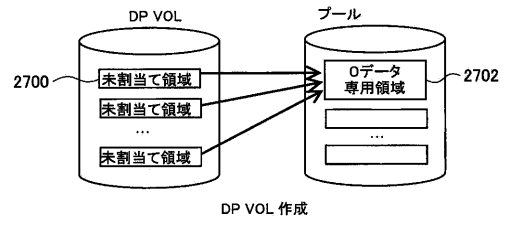
【図26】

図26



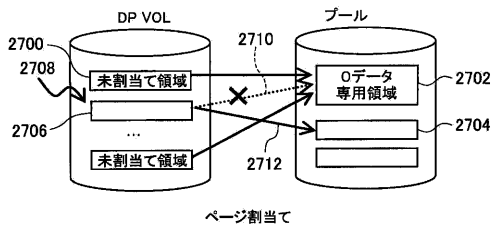
【図27A】

図27A



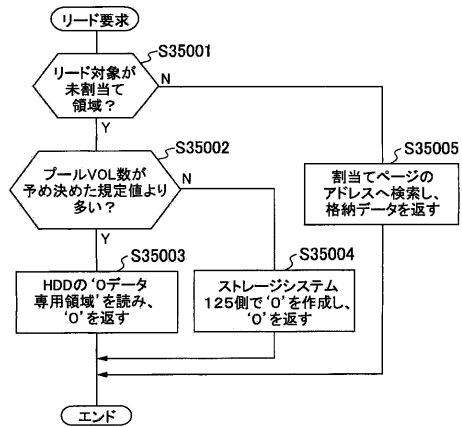
【図27B】

図27B



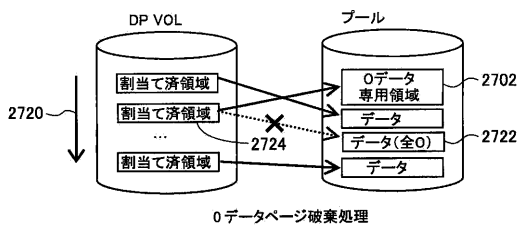
【図28】

図28



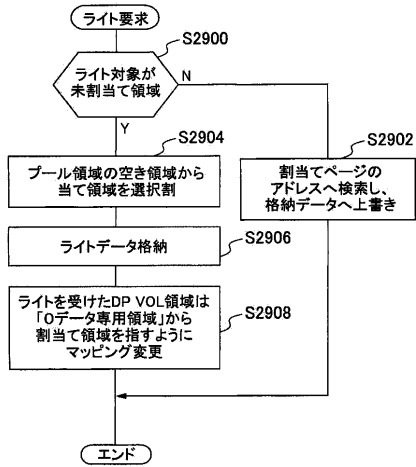
【図27C】

図27C



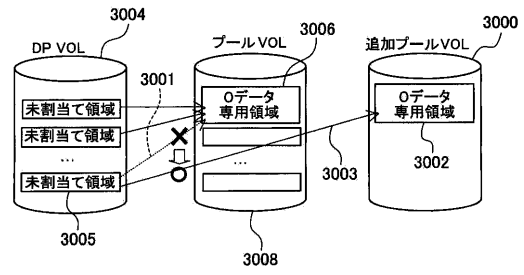
【図29】

図29



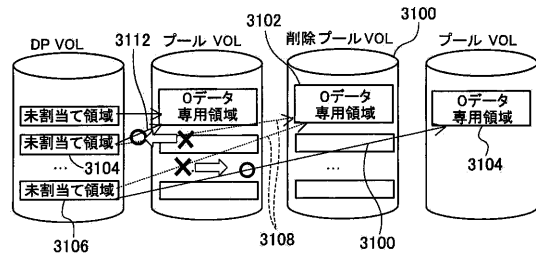
【図30A】

図30A



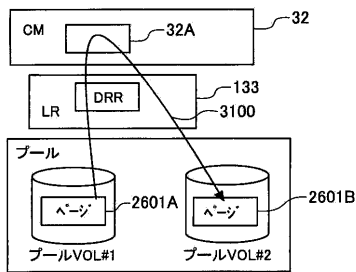
【図30B】

図30B



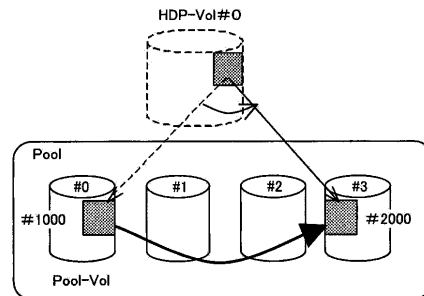
【図31】

図31



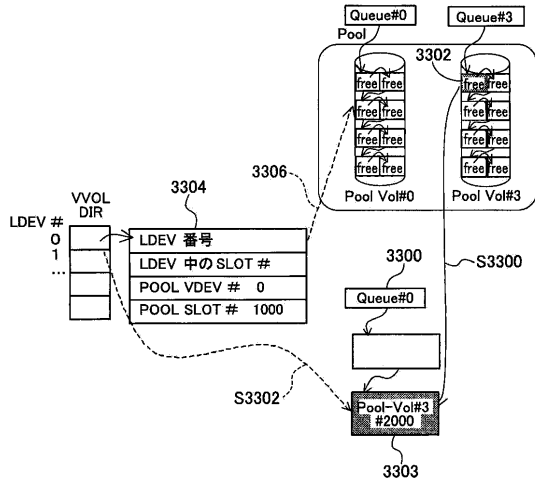
【図32】

図32



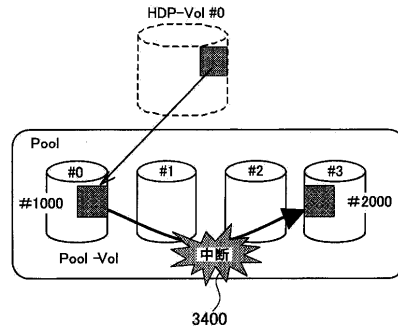
【 図 3 3 】

図33



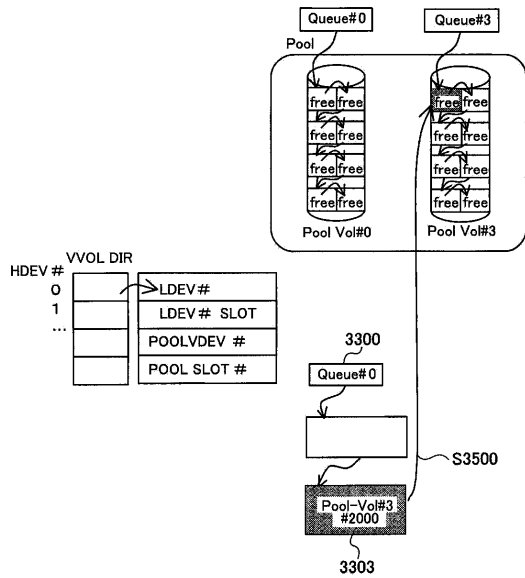
【 図 3 4 】

図34



【 図 3 5 】

図35



【 図 3 6 】

図36

マイグレーション進捗管理テーブル

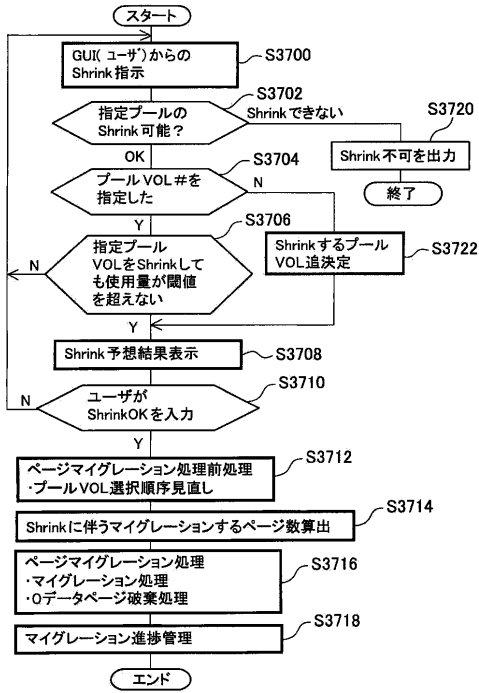
| プール# | 移動対象 HDP VOL数 | 移動済み HDP VOL数 | プールの状態 |
|------|---------------|---------------|--------|
|      |               |               |        |

| プールVOL# | 移動対象 ページ数 | 移動残 ページ数 |  |
|---------|-----------|----------|--|
|         |           |          |  |

【 図 3 7 】

図37



【 図 3 8 】

図38

| プール# | プールの状態   | プールの容量 (GB) | プールの使用量 (GB) | プールの空き容量 (GB) |
|------|----------|-------------|--------------|---------------|
| 0    | Shrink 中 | × × ×       | × × ×        | × × ×         |
| 1    | Shrink 未 |             |              |               |
| 2    |          |             |              |               |

プール#0

| プールVOL # | プールのVOL 状態 | プールのVOL 容量 (GB) | プールのVOL 使用量 (GB) | プールのVOL 空き容量 (GB) |
|----------|------------|-----------------|------------------|-------------------|
| 0        | 正常         | × × ×           | × × ×            | × × ×             |
| 1        | 閉塞         |                 |                  |                   |

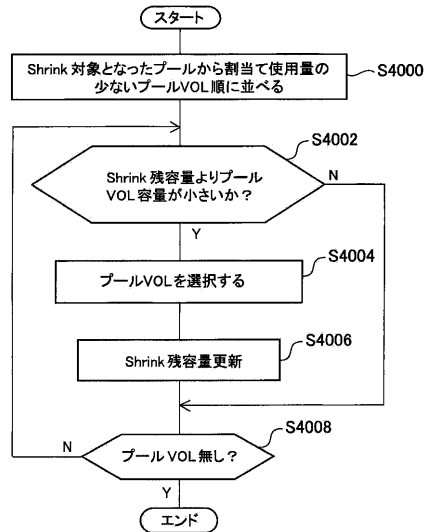
【 図 3 9 】

図39

| プール # | Shrink 前  |            |             |          | Shrink 後(予想) |            |             |          |
|-------|-----------|------------|-------------|----------|--------------|------------|-------------|----------|
|       | プール容量(GB) | プール使用量(GB) | プール空き容量(GB) | 空き割合 (%) | プール容量(GB)    | プール使用量(GB) | プール空き容量(GB) | 空き割合 (%) |
| 0     | × × ×     | × × ×      | × × ×       |          | × × ×        | × × ×      | × × ×       |          |

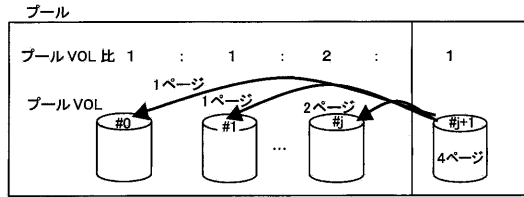
【 図 4 0 】

図40



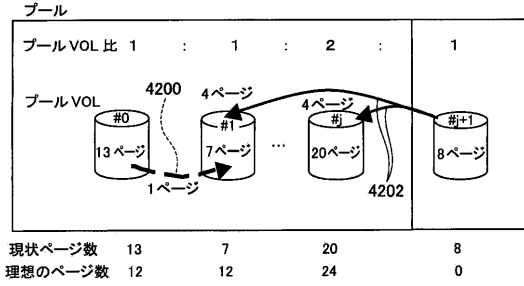
【図41】

図41



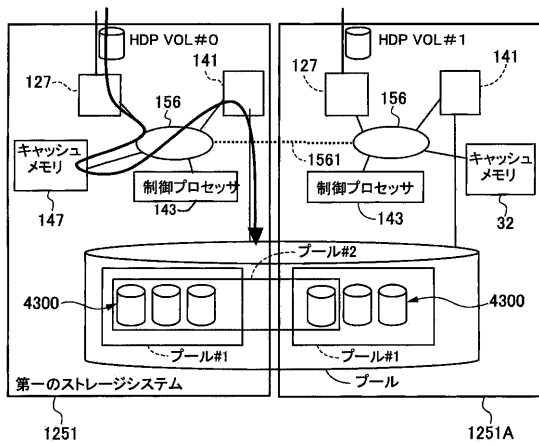
【図42】

図42



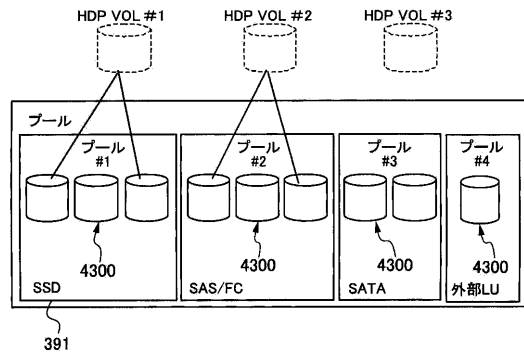
【図43】

図43



【図44】

図44



---

フロントページの続き

合議体

審判長 小曳 満昭

審判官 白石 圭吾

審判官 山田 正文

- (56)参考文献 特開2003-234158(JP,A)  
特開2007-79885(JP,A)  
特開2009-48514(JP,A)

- (58)調査した分野(Int.Cl., DB名)  
G06F 3/06 - 3/08